

**AN ELECTROPHYSIOLOGICAL INVESTIGATION
OF REWARD PREDICTION ERRORS IN THE
HUMAN BRAIN**

THOMAS DANIEL SAMBROOK

This copy of the thesis has been supplied on condition that anyone who consults it is understood to recognise that its copyright rests with its author and that no quotation from the thesis and no information derived from it may be published without the author's prior consent.

**AN ELECTROPHYSIOLOGICAL INVESTIGATION
OF REWARD PREDICTION ERRORS IN THE
HUMAN BRAIN**

by

THOMAS DANIEL SAMBROOK

A thesis submitted to Plymouth University
in partial fulfilment for the degree of

DOCTOR OF PHILOSOPHY

School of Psychology

March 2015

AN ELECTROPHYSIOLOGICAL INVESTIGATION OF REWARD PREDICTION ERRORS IN THE HUMAN BRAIN

THOMAS DANIEL SAMBROOK

Reward prediction errors are quantitative signed terms that express the difference between the value of an obtained outcome and the expected value that was placed on it prior to its receipt. Positive reward prediction errors constitute reward, negative reward prediction errors constitute punishment. Reward prediction errors have been shown to be powerful drivers of reinforcement learning in formal models and there is thus a strong reason to believe they are used in the brain. Isolating such neural signals stands to help elucidate how reinforcement learning is implemented in the brain, and may ultimately shed light on individual differences, psychopathologies of reward such as addiction and depression, and the apparently non-normative behaviour under risk described by behavioural economics. In the present thesis, I used the event related potential technique to isolate and study electrophysiological components whose behaviour resembled reward prediction errors. I demonstrated that a candidate component, “feedback related negativity”, occurring 250 to 350 ms after receipt of reward or punishment, showed such behaviour. A meta-analysis of the existing literature on this component, using a novel technique of “great grand averaging”, supported this view. The component showed marked asymmetries however, being more responsive to reward than punishment and more responsive to appetitive rather than aversive outcomes. I also used novel data-driven techniques to examine activity outside the temporal interval associated with the feedback related negativity. This revealed a later component responding solely to punishments incurred in a Pavlovian learning task. It also revealed numerous salience-encoding components which were sensitive to a prediction error’s size but not its sign.

Table of Contents

0 Contents

1	General Introduction	1
1.1	The Reinforcement Learning Theory of the ERN	1
1.2	Scope of the Thesis	4
1.3	Is the FRN a Positive RPE, a Negative RPE or Both?	4
1.4	Is the FRN Insensitive to Expected Value?	6
1.5	Is the FRN Insensitive to Whether RPE Size Is Modulated by the Magnitude of the Outcome or its Likelihood?	8
1.6	Is the RPE Indexed by the FRN a General Valuation or Specific to a Reinforcement Learning Module?	9
1.7	Where and When Does the FRN Occur?	11
1.8	Glossary	12
2	Prospect Theory Does Not Describe the Feedback Related Negativity Value Function	16
2.1	Chapter Abstract	16
2.2	Introduction	16
2.3	Experiment 1	22
2.3.1	Method	22
2.3.2	Results	28
2.3.3	Discussion	38
2.4	Experiment 2	38
2.4.1	Method	38
2.4.2	Results	39
2.4.3	Discussion	49
2.5	General Discussion	49
3	An Examination Of Sensitivity To Positive, Negative And Unsigned Prediction Errors	55
3.1	Chapter Abstract	55
3.2	Introduction	55
3.3	Methods	63
3.3.1	Participants	63
3.3.2	Task Rationale	63
3.3.3	Task Procedure	64
3.3.4	EEG Recording	66
3.3.5	EEG Analysis	67
3.4	Results	70
3.4.1	Behavioural Data	70

3.4.2	Electrophysiological Data	72
3.4.3	Discrimination of Salience and Utility effects.	79
3.4.4	Domain Effects.....	83
3.5	Discussion	85
4	An Examination of Sensitivity to Positive, Negative and Unsigned Prediction Errors with Principal Components Analysis	94
4.1	Chapter Abstract.....	94
4.2	Introduction	95
4.3	Methods	103
4.3.1	Participants	103
4.3.2	Task	104
4.3.3	Modelling of RPEs.....	106
4.3.4	EEG Recording	107
4.3.5	EEG and PCA Analysis	108
4.4	Results	110
4.4.1	Behavioural Results	110
4.4.2	ERPs.....	111
4.4.3	PCA Components.....	115
4.5	Discussion	121
5	A Meta-Analysis of the FRN Using Great Grand Averages.....	127
5.1	Chapter Abstract.....	127
5.2	Introduction	127
5.2.1	A Meta-Analysis of the FRN	127
5.2.2	Existing Evidence for Modulation of the FRN by Magnitude and Likelihood	128
5.2.3	Problems with FRN Measurement and Implications for Meta-analysis .	129
5.2.4	A Novel Method: Great Grand Averages	131
5.2.5	Moderators	133
5.3	Methods	140
5.3.1	Inclusion and Exclusion Criteria.....	140
5.3.2	Moderator Analyses	142
5.3.3	Search Strategies	143
5.3.4	Coding Procedures: Generating Great Grand Averages Waveforms.....	145
5.3.5	Coding Procedures: Quantifying the FRN	146
5.3.6	Statistical Methods	151
5.4	Results	156
5.4.1	Modulation of the FRN by Magnitude and Likelihood.	156
5.4.2	Publication Bias	163
5.4.3	Moderator Analyses of the FRN	165

5.4.4	Validation of the GGA Technique	168
5.4.5	Main Effects of Magnitude and Likelihood	176
5.5	Discussion	179
5.5.1	The RPE-FRN and Main Effects of Valence, Magnitude and Likelihood 179	
5.5.2	Applications of the Present Findings	182
5.5.3	Implications for the Measurement of the RPE-FRN and FRN	186
5.5.4	Evaluation of the GGA Technique.....	188
6	General Discussion.....	191
6.1	The Reinforcement Learning Theory of the Error Related Negativity	191
6.2	Is the FRN a Response to +RPEs, –RPEs or Both?	192
6.3	Is the FRN Insensitive to Expected Value?.....	193
6.4	Is the FRN Indifferent to Whether RPE Size is Modulated by the Magnitude of the Outcome or its Likelihood?	194
6.5	Where and when does the FRN occur?	195
6.6	Is the RPE Indexed by the FRN a General Valuation or Specific to a Reinforcement Learning Module?	197
6.7	How should the FRN be studied?	198
7	References	199

Figures

Figure 1 The prospect theory value function	21
Figure 2 Sequence of stimuli in Experiment 1	24
Figure 3 Participants' affective ratings in Experiment 1	29
Figure 4 The effect of feedback valence in Experiment 1	31
Figure 5 FRN in Experiment 1	33
Figure 6 Difference waves in Experiment 1	35
Figure 7 P3 in Experiment 1	37
Figure 8 Participants' affective ratings in Experiment 2	41
Figure 9 Feedback valence in Experiment 2	43
Figure 10 FRN in Experiment 2	45
Figure 11 Difference waves in Experiment 2	46
Figure 12 P3 in Experiment 2	48
Figure 13 A schematic representation of possible response functions to RPE utility.	57
Figure 14 Summary of a single trial of the task procedure	65
Figure 15 Effects of valence	73
Figure 16 Correlational waveforms	75
Figure 17 A factorial rendition of the experimental design	78
Figure 18 Grand average of voltages at each level of RPE utility	82
Figure 19 Correlational waveforms broken down by outcome domain.....	84
Figure 20 How component overlap undermines simple effects analysis of the FRN.....	98
Figure 21 Summary of a single trial of the task procedure	104
Figure 22 Grand average ERPs in response to +RPEs vs -RPEs	112
Figure 23 Correlational waveform prior to PCA	114
Figure 24 Factors extracted under PCA	116
Figure 25 Factors extracted by the PCA but not selected for further analysis.....	117
Figure 26 Factor TF3/SF1	119
Figure 27 How the FRN is studied.....	149
Figure 28 Simple GGA waveforms.....	158
Figure 29 FRN GGAs showing effect of RPE-FRN	158
Figure 30 Forest plot showing RPE-FRN simple effect size	161
Figure 31 Funnel plots for the unweighted simple effect size of the RPE-FRN	164
Figure 32 Moderator analysis of the FRN.....	167
Figure 33 Main effects of magnitude, likelihood and valence (RPE-FRN also)	178

Tables

Table 1 Properties of feedback under different economic dimensions	68
Table 2 Factor combinations selected for statistical analysis	115
Table 3 Experiments used in the meta-analysis	135
Table 4 Results of a conventional vs. GGA meta-analysis	171
Table 5 Effect sizes in the validation data set	172
Table 6 Effect sizes of the RPE-FRN calculated from published data	175

Acknowledgements

I would like to thank my supervisor, Dr. Jeremy Goslin, for his many and diverse contributions to this body of work.

AUTHOR'S DECLARATION

At no time during the registration for the degree of Doctor of Philosophy has the author been registered for any other University award without prior agreement of the Graduate Committee.

Work submitted for this research degree at the Plymouth University has not formed part of any other degree either at Plymouth University or at another establishment

A programme of advanced study was undertaken. Relevant scientific seminars and conferences were regularly attended at which work was often presented and several papers prepared for publication.

Publications:

Sambrook, T. D., & Goslin, J. (revise and resubmit). Principal components analysis of reward prediction errors in a reinforcement learning task. *NeuroImage*.

Sambrook, T. D., & Goslin, J. (2015). A neural reward prediction error revealed by a meta-analysis of ERPs using great grand averages. *Psychological Bulletin* 114 213-235.

Sambrook, T. D., & Goslin, J. (2014). Mediofrontal event-related potentials in response to positive, negative and unsigned prediction errors. *Neuropsychologia*, 61 1-10.

Sambrook, T. D., Roser, M., & Goslin, J. (2012). Prospect theory does not describe the feedback-related negativity value function. *Psychophysiology*, 49(12), 1533-1544.

Presentation and Conferences Attended:

Sambrook, T. D., & Goslin, J. (2014). *Time of the signs: Event-related potentials code unsigned and positive reward prediction errors in different intervals*. Paper presented at symposium: "EEG correlates of reward and prediction errors: a biting challenge", British Association for Cognitive Neuroscience, Sep 11th 2014, York, United Kingdom.

Sambrook, T. D., & Goslin, J. (2014). *A reward prediction error is revealed by ERP great grand averages*. Paper presented at the winter meeting of the Experimental Psychology Society, London, Jan 9th 2014 .

Word count of main body of thesis: 49,541

Signed.....

Date.....

1 General Introduction

1.1 The Reinforcement Learning Theory of the ERN

In 2002, Holroyd and Coles published an influential paper claiming that EEG could be used to observe a learning signal in the human brain (Holroyd & Coles, 2002). The theory integrated two separate event related potential (ERPs) components. The first of these, known as the error related negativity (ERN), arose when participants made motor errors in speeded tasks, and appeared as a negativity approximately 100 ms after error commission. For the ERN to occur, the participant had to be aware that they had made an error. The second component, which has come to be known as the feedback related negativity (FRN), arose when participants performed a task in which knowledge of whether an error had been committed was not fully knowable at the time the participant made a response, perhaps because there was some probabilistic basis for what the correct response was on a given trial. In this case, the negativity was observed not after the participants response but after externally communicated feedback, appearing with a latency of ~300 ms.

The authors claimed that these signals had a common basis, both functionally and anatomically. Functionally, they claimed that in each case the signal reflected a reward prediction error (RPE), that is the size of the difference between an expected reward and an obtained reward for a particular trial, with errors indicating that an expected reward would not be obtained on the present trial. The following example would typify the ERN. A participant has acquired an expectation that pressing a "Z" in response to a certain stimulus always gives a reward of 10p, but pressing "M" always gives no reward. When, on a given trial they are provided with an opportunity to make this response, they generate an expectation of 10p: this is the expected value (EV) of

this trial. If they then incorrectly press M (typically because the task is speeded), they internally register this error and revise their expectation of reward to expectation of no reward. This revision takes the form of a RPE, necessarily negative in the present example, with its size determined by the amount of forfeited reward, so here -10p.

In the case of the FRN similar principles apply, but the expected value of the reward is somewhat different. Here, responses are only probabilistically related to rewards. To best contrast with the ERN we take the most extreme case in which “Z” and “M” each return a reward of 10p fifty percent of the time, and randomly. In this case, the expected value of pressing either key is its average return which is 5p. If participants press the wrong key, the error only comes to light when they are provided with external feedback. The prediction error will be the size of the forfeited expected value, in this case being -5p.

In both tasks, expected value, and therefore the resulting RPE, can be varied, and Holroyd and Coles claimed that the size of the negative deflection shown by the ERN and FRN was related to the size of the prediction error. While the tasks that elicit the two ERPs may seem rather different, Holroyd and Coles’ insight was that both kinds of tasks could produce sudden downward revisions in reward expectation which might harness the same functional system. This was a reinforcement learning system devoted to improving behavioural performance by tracking errors. Their claim that these errors were instantiated as RPEs, that is, as actual scalar quantities of lost value, rather than merely categorically as “errors”, builds on the very powerful role that these terms have been shown to hold in reinforcement learning (e.g. Sutton & Barto 1998). Indeed, given their simplicity and economy, we would be surprised if human reinforcement learning did not use RPEs. Holroyd and Coles also argued that the ERN and FRN belonged to a specific class of RPE known as temporal prediction errors. Temporal prediction errors are generated not only when an expected reward is missed, but whenever the

expectation of reward changes, thus in an experimental context, they might occur at multiple points during a trial prior to final reward delivery or omission. Temporal prediction errors allow both an “in-running” revision of expected value, and also the generation of error signals to correct whatever behaviour led to the reduction of EV. This provides a more sensitive basis for learning than simply computing an error signal at the time of reward omission and then applying this to all preceding behaviours that led to the omission.

Holroyd and Coles claimed that a common neural substrate underlay both the FRN and ERN. They argued that these ERPs reflected the arrival of a prediction error generated by the midbrain dopamine system at the anterior cingulate cortex (ACC). The ACC is highly plausible in this regard, in that it is widely believed to mediate “conflict” (a class of event to which prediction errors belong), and also because it is known to have a role in action selection. It would thus be ideally placed to use prediction errors to guide learning of optimal behaviour. In turn, the claim that the prediction error has its origin in the midbrain dopamine system is strongly supported, with single cell work convincingly showing that prediction errors in single cells are coded by change in phasic firing (Schultz, Dayan, & Montague, 1997).

Holroyd and Coles’ theory entails a number of claims about the FRN: that it is functionally related to the ERN, that it is dopaminergically generated, that it arises in the ACC, that it is limited to negative RPEs, and that it forms part of a reinforcement learning system. FRN studies are numerous however, and are by no means limited to testing these claims, nor necessarily assume that all of them must be true. One particularly apparent shift in emphasis is a move towards viewing the FRN as a reward valuation. Holroyd and Coles’ emphasis, and that of Miltner, Braun, and Coles (1997), who first identified the component, was on error commission. The relevance of reward simply arose because Holroyd and Coles argued that that it was a reinforcement

learning system that produced this component, and thus reward had to be the quantity coded. The FRN literature has over time drifted away from characterising the FRN as an error signal, and instead tended to view it as a valuation, retaining the error aspect only insofar as it is a reward prediction “error”. In fact, a contemporaneous paper argued that when reward value and error commission were dissociated, the component was better described by its being a valuation (Gehring & Willoughby, 2002).

1.2 Scope of the Thesis

Following the publication of Holroyd and Coles’ theory, competing claims have been made for what the FRN codes, including error commission, RPEs, salience, affective value, equivocality of feedback. Because the theories in many cases make very similar predictions, an exhaustive testing of them lies beyond the scope of this work, although circumstantial evidence will emerge that is inconsistent with the claims of some theories. In this thesis we make the provisional assumption that the FRN does indeed code a RPE and set out to characterise the nature of this RPE. While the predictive power of the RPE account is not explicitly compared to that of alternative accounts, nevertheless we argue that the work in this thesis promotes the credibility of the RPE account. The findings are fully consistent with that account, and in some cases contradict rival theories such as the salience account. In making this case, the following key questions are addressed.

1.3 Is the FRN a Positive RPE, a Negative RPE or Both?

Insofar as outcomes can be better than expected as well as worse, RPEs can be positive (+RPEs) as well as negative (–RPEs). The most powerful formal reinforcement learning models make use of both terms in optimising learning. Formally, the sign of

the RPE denotes whether an outcome constitutes a reward or punishment, allowing the possibility of a synthesis of computational models of reinforcement learning with the wider behavioural literature on the differential effects of reward and punishment. Thus, establishing the relative sensitivity of the FRN to +RPEs and –RPEs, is an important step in discovering its role in learning and behaviour.

In Holroyd and Coles' theory the FRN was conceived of as a –RPE. This was necessary in order to make the case that it was functionally coupled with the ERN, for which no +RPE encoding can occur. It is also possible that this assumption arose because of the visually compelling negative deflection shown on trials in which rewards were missed, which led to the component's name.

Subsequent studies have questioned the assumption that only –RPEs are encoded by the FRN. First, peaks observed on individual waveforms are now recognised as unreliable indicators of underlying components because they are the sum of many overlying components (Luck, 2005). It is thus possible that a component quite unrelated to reward processing produces a negative deflection at the latency of the FRN under all trials of an experiment. As such, the negative deflection observed under reward omission is indicative of *absence* of any reward related activity, while the removal of this peak in the case of rewards is indicative of a superimposed component coding a +RPE. Such a conjecture has indeed been made by Holroyd himself, who has proposed changing the component's name to the "feedback correct-related positivity" (Holroyd, Pakzad-Vaezi, & Krigolson, 2008).

Second, the FRN is proposed to have its origin in the phasic firing of dopaminergic neurons of the midbrain, with the original account suggesting that reward omission produced a dip in the phasic firing of these neurons, constituting the –RPE sent to the ACC. However, while such dips do occur in dopamine neurons, because of the low tonic level of firing, the dip is small compared to the rise in firing observed to

follow +RPEs. If the midbrain dopamine system is the origin of the RPE that goes on to produce the FRN, then the FRN should be more responsive to +RPEs than –RPEs, or at least show better discriminability between the size of +RPEs, because of the increased range of dopamine release associated with them.

This question of the sensitivity of the FRN to +RPEs and –RPE has been regularly visited in the literature, but the treatment has not generally been very sound. A common strategy has been to simply compare the difference in voltage of the response to small and large +RPEs to that of small and large –RPEs, and to use this to infer any selective sensitivity. However, the likely presence of overlying components renders this methodologically unsafe. Other studies which have compared the size of +RPEs and –RPE to a neutral baseline of breaking even suffer from the same problem. Chapters 3 and 4 address the question of whether the FRN responds to both +RPEs and –RPEs, and in those chapters, removal of the effects of overlying components is a central concern.

1.4 Is the FRN Insensitive to Expected Value?

Formally, an RPE is conceived of merely as the signed difference between an expected and obtained value. The actual expected value that was used to obtain this difference is irrelevant. If the FRN reflects such an RPE then it should be indifferent to whether the expected value was, say a loss or gain of money, or indeed the size of the expected loss or gain. Generalising to more biologically relevant stimuli, if the FRN simply codes for how much better or worse than expected an outcome is, then its response to acquiring a larger than expected food item should be similar to its response to a smaller than expected bout of pain. Of course in this latter case, the modality (food vs. pain) also differs and these may involve radically different valuation systems. However, the general important point is that formal reinforcement learning systems do not represent whether either an expected or obtained outcome is in some sense aversive

(pain, monetary deduction) or appetitive (food, monetary increment), merely the sign and the size of the difference of expected and obtained outcomes. This derived term, the RPE is then represented on a single unidirectional scale running from very bad to very good.

The question is of interest for a number of reasons. One is simply that it needs to be answered in order to fully characterise the FRN's role in neural valuation. A second is that it constitutes a test case of the degree to which formal AI models can be expected to reflect the actual solutions that have evolved: perhaps appetitive outcomes such as food and aversive outcomes such as food have experienced quite separate selection pressures and developed in different neural systems. A third is methodological: while all FRN experiments provide controlled variation of the valence of the outcome with respect to expected value (i.e. produce better than or worse than expected outcomes) they rarely control the outcomes' inherent value as described above. This produces consistent confounds, which, as we describe in Chapter 3, might lead to an apparent sensitivity to +RPEs that in fact has its origin in a greater sensitivity to inherently appetitive outcomes. A final reason why the question is of interest relates to prospect theory (Kahneman and Tversky 1979), a hugely influential theory of human valuation which accounts for a number of suboptimal aspects of human behaviour. This claims that rational decision-making, which should solely be guided by increasing expected value, is in practice affected by the actual expected values involved. This produces the bias of loss aversion and an inappropriate sensitivity to marginal utility. In Chapter 3 we test whether the FRN shows these prospect theoretic biases or whether, in contrast, it fits better to "EV indifferent" formal reinforcement learning models.

1.5 Is the FRN Insensitive to Whether RPE Size Is Modulated by the Magnitude of the Outcome or its Likelihood?

When a +RPE is incurred, this may be because the obtained reward was of an expected magnitude but its prior likelihood of actually being obtained was not certain. Alternatively, it may be that a reward was expected, or even certain, but the reward's magnitude was high compared to the distribution of magnitudes available.

Experimentally, RPEs can thus be modulated by varying either the likelihood or magnitude of rewards (or indeed punishments). Of course, in real ecologies, prediction errors are likely to result from variations of both likelihood and magnitude of reward, however the dichotomy described above usefully isolates the two sources of potential RPE modulation. Furthermore, it accurately describes the literature insofar as most FRN experiments use one means of modulation or the other. That is, they either offer a fixed reward whose likelihood of receipt varies, or they offer a range of reward magnitudes, each of which is as likely as the other. Both schemes will produce +RPEs and -RPEs of differing sizes whose effect on the FRN can then be studied.

The question then arises whether the FRN is indifferent to the source of RPE modulation. Formal reinforcement learning models are indifferent, the size and sign of the RPE is all that matters, the historical details of how it arose are not retained. The dopamine responses of the midbrain appear to respect this, with RPEs of similar size producing similar changes of phasic firing regardless of the modulator (Schultz, 2010). Since this is the signal believed to underlie the FRN, it should likewise be indifferent.

However a highly influential paper by Yeung and Sanfey (2004) has claimed that the FRN is insensitive to modulation of RPE size by reward magnitude, and a number of following papers have echoed this claim. This has theoretical ramifications for the FRN. Since we can assume that humans are in reality sensitive to the magnitude of better and worse than expected outcomes, the claim implies that the FRN can only

form part of a wider apparatus of reinforcement learning. While the monolithic treatment of RPEs in formal models is efficient and powerful, it need not be reflected in the evolved mechanisms of the brain. In real ecologies some classes of natural “goods” are likely to continuously vary in magnitude (the quantity of food in a discovered food patch, say) while others are essentially binary but vary in likelihood (the availability of a mating opportunity, or the danger of a predator, say). Specialist neural systems devoted to these ecological classes may thus have evolved their own separate RPE codings that are insensitive to either reward magnitude or likelihood.

An alternative possibility is that the FRN does indeed perform a generalised RPE encoding, incorporating both likelihood and magnitude modulated RPEs, but suffers component overlap with components registering magnitude and surprise as psychological rather than economic properties, which distort the overall picture, possibly cancelling out an RPE signal. For this reason, in Chapter 4 we examine responses to magnitude and likelihood RPEs separately, performing PCA to extract the various components.

A further possibility is that the FRN is indeed responsive to both magnitude and likelihood but that variability in measurement of the FRN, along with a tacit expectation of non-sensitivity to magnitude following Yeung and Sanfey’s (2004) paper has obscured the true picture. We investigate this with a meta-analysis in Chapter 5.

1.6 Is the RPE Indexed by the FRN a General Valuation or Specific to a Reinforcement Learning Module?

Balleine, Daw, and O’Doherty (2008) have proposed that the brain computes three kinds of expected values, Pavlovian, habit and goal values. Pavlovian values denote the expected value of a stimulus. Habit values and goal values denote the expected value of an action performed on that stimulus. In the case of habit values, this

expected value is simply the cached average of reinforcement experienced from previously performing that action, and habit values thus constitute basic terms in instrumental conditioning. In the case of goal values, the expected value is derived from a model of the environment, built partly from cached reinforcement but also from wider propositional reasoning, which can incorporate knowledge of higher order relationships, allowing, for example, rapid behavioural adjustment in reversal learning tasks, which model free reinforcement learning is much slower to respond to.

If the FRN is an RPE, the question must be asked: against which expected value term is its value computed? While the single cell RPEs that supposedly underlie the FRN are near universally studied with Pavlovian conditioning, FRN tasks are most usually presented as instrumental conditioning tasks, in which participants must optimise a response. Indeed in its original formulation, the RL-ERN necessarily entailed an instrumental conditioning role for the FRN insofar as the model made use of eligibility traces. That is to say the model assumed that the motor controller executing an action retained an eligibility trace for a short period after acting, which rendered it capable of learning from a resulting RPE. Support for the role of eligibility traces in generating the FRN was recently provided by Weinberg, Luhmann, Bress, and Hajcak (2012), who showed that delaying feedback to 6 s after action removed the FRN. Furthermore, Yeung, Holroyd, and Cohen (2005) showed a greatly reduced FRN when participants were told they could not control outcomes, effectively changing the task from an instrumental to Pavlovian conditioning task. Nevertheless, other factors, such as task interest, were shown to be mediating variables, and other researchers (e.g. Talmi, Atkinson, & El-Deredy, 2013) have found FRNs to Pavlovian tasks. Chapters 3 and 4 present very similar tasks in Pavlovian and instrumental formats, to allow a better comparison of their respective effects on the FRN.

Whether the FRN is computed against an EV term derived simply from reinforcement history (habit values) or whether it incorporates wider knowledge in order to produce goal values, is a question that is only recently receiving attention. Walsh and Anderson (2011b) showed that the FRN was insensitive to instructional information and dependent on reinforcement history, suggesting it was computed against a habit value, with Chase, Swainson, Durham, Benham, and Cools (2011) providing supporting evidence. We go some way to addressing this question by comparing the FRN's response to the verbal, affective ratings given by participants, which can be assumed to be based on wider cognition.

1.7 Where and When Does the FRN Occur?

This question is of both theoretical interest in tying the FRN to wider brain function, but also of pressing methodological importance, certainly in the case of the component's temporal profile, since shifting the window of EEG measurement by even a few tens of milliseconds can radically change statistical effects. Identifying the true latency of the FRN has been hampered by inconsistencies in what the term refers to. In some cases it is used to refer to an observed negative deflection, present for negative outcomes but absent or reduced for positive ones. In other cases it is used to refer to the experimental effect of valence, typically expressed as a negative peaked difference wave constructed by subtracting good from bad outcomes. There is no necessary reason why these two negativities should peak at the same latency. Furthermore, in both cases, there is great variability in the interval in which measurement is made, and one of the objectives of the meta-analysis of the FRN presented in Chapter 5 was to harness "the wisdom of crowds" to establish where the FRN typically occurred and thus where it should be measured in future experiments.

An additional complication, and one pervasive in electroencephalography, is that while the term FRN is sometimes used to refer to the observed behaviour of scalp recorded grand average ERPs, it is also used to refer to the underlying neural generator of these potentials. Because of component overlap, the former only approximates to the latter. Thus the observed FRN may in fact be a composite of more than one valence sensitive component. To assess this, in Chapter 4 we use PCA to attempt to separate out the components that make up the scalp grand averages and so better approximate the underlying neural components.

While originally localised to the ACC by Holroyd and Coles, a number of recent papers have applied source localisation to the FRN component after its extraction using PCA, and identified the basal ganglia as its source. While source localisation of the FRN is not the principal focus of this thesis, which instead focusses on the FRN's functional response to RPEs, in Chapter 4 we nevertheless perform source localisation on the FRN after its extraction by PCA.

1.8 Glossary

Elucidation of the factors to which the FRN is sensitive has not been helped by inconsistencies in the terminology that the literature uses to describe the basic dimensions along which valuation is made. The glossary given below sets out the usage that will apply for this thesis. We do not prescribe this for the field overall, it is simply provided as a reference for this thesis.

Valence

This is a categorical binary variable that describes whether an outcome is better or worse than its prior expected value. It is thus used exclusively to refer to RPEs and does not refer to any inherent goodness or badness of an outcome. Its levels are good vs.

bad. Where the context is appropriate we also refer to these levels as positive RPEs vs. negative RPEs, +RPEs vs. -RPEs or reward vs. punishment. Note that in this regard we use the terms reward and punishment to refer to internal events rather than as manifestations of behaviour: a neuroeconomic rather than behavioural approach

Outcome Value

The face value of the outcome: if the computer displays +10p then this is the outcome value.

Domain

This is a categorical variable describing where an outcome lies with respect to some wider reference point which is typically *not* EV and which serves to mark outcomes on either side of it as inherently good and others as inherently bad. It has two levels when referring to outcomes: gain and loss. It is equivalent to the distinction between appetitive and aversive outcomes by some authors usage of these terms, though other authors have used those terms synonymously with reward and punishment. The reference point will often be unambiguous. For thirsty participants, all deliveries of water are gain domain outcomes even if they smaller than expected. For normal participants, all electric shocks are loss domain outcomes, even if they are milder than expected. For the monetary outcomes used in this thesis, loss outcomes are deductions from the monetary holding the participant currently enjoys, gain outcomes are increases. When referring to prospects, domain has three levels: gain where all possible outcomes will result in an increase in monetary holdings, or at least no loss; loss, where all possible outcomes will result in a decrease in monetary holdings, or at least no gain; and mixed, where both losses and gains are possible.

Expected Value (EV)

The sum of all potentially available outcomes weighted by their likelihood. EVs are signed numbers.

Reward Prediction Error, RPE

The difference between an outcome and its expected value. RPEs are signed numbers.

Utility

The dimension over which RPEs vary, running from $-\infty$ to $+\infty$.

Unsigned Prediction Error, UPE

The absolute difference between an outcome and its expected value. UPEs are unsigned numbers.

Salience

The dimension over which UPEs vary, running from 0 to $+\infty$.

Feedback Related Negativity, FRN

A valence encoding ERP component revealed by a negative peaked difference wave constructed by subtracting the voltage produced by good outcomes from that produced by bad outcomes.

RPE-FRN

A (putative) reward prediction error encoding ERP component revealed by a negative peaked difference wave constructed by subtracting the FRN for large good and bad outcomes from that for small good and bad outcomes.

Great Grand Average

An average of a number of grand averages taken from different experiments, used to provide a graphic demonstration of the results of a meta-analysis.

2 Prospect Theory Does Not Describe the Feedback Related Negativity Value Function

The work in this chapter is strongly based on a published paper (Sambrook et al., 2012)

2.1 Chapter Abstract

Prospect theory accounts for non-optimal economic behaviour by positing a value function in which possible losses are over-weighted compared to possible gains and the marginal utility of rewards decreases with size. We investigated whether the prospect theoretic effects shown in behavioural and fMRI studies were present in the FRN. Participants' stated satisfaction with outcomes followed prospect theory but their FRNs did not, instead showing no effect of marginal utility and greater sensitivity to potential gains than losses.

2.2 Introduction

Making effective decisions under choice requires people to construct mental representations of *value*. When goods on offer will be received with a certain probability only, a mathematically correct approach is to base decisions on expected value (EV), that is, the product of a good's value and the probability of obtaining it. Humans typically depart from this solution and the science of behavioural economics attempts to account for this. The most successful account to date is prospect theory (Kahneman & Tversky, 1979; Tversky & Kahneman, 1992). Prospect theory identifies three ways in which observed choices typically deviate from those dictated by expected value. First, following von Neumann and Morgenstern (1944), choices are based on the

subjective value of the good. Subjective value shows diminishing growth with respect to objective value, producing the phenomenon of diminishing marginal utility. Second, this utility function is asymmetrically generated about a mathematically arbitrary but psychologically salient anchor point. The anchor point is typically zero, the utility function is typically steeper for outcomes less than zero, and this gives rise to the phenomenon of loss aversion, namely that losses hurt more than equivalent gains please. Third, probabilities are distorted at the decision stage, such that small probabilities are inflated and large ones are diminished. It is the first two of the distortions described above, those that reify the zero anchor point, and are collectively known as the “value function”, that are the subject of the current paper.

fMRI studies have identified a wide network of areas that appear responsive to value and there is growing evidence that some of these may show a value function that is anchor point dependent. Pine et al. (2009) found that BOLD activation in the striatum, anterior cingulate cortex (ACC) and ventral tegmental area showed effects consistent with diminishing marginal utility. Loss aversion has been studied by Tom, Fox, Trepel, and Poldrack (2007), who found that the striatum and ventromedial prefrontal cortex showed greater decreases in activity following losses than increases following gains, and that this effect correlated with individual behavioural differences in loss aversion. Loss aversion brain-behaviour correlates have also been shown by De Martino, Kumaran, Holt, and Dolan (2009) and Seymour, Daw, Dayan, Singer, and Dolan (2007), specifically in the ventral striatum. In contrast, other studies have identified the amygdala and insula as sources of loss aversion (De Martino, Kumaran, Seymour, & Dolan, 2006; Knutson et al., 2008; Weber et al., 2007). The wider evidence for prospect theory effects in the brain is reviewed in Trepel, Fox, and Poldrack (2005).

A quite different body of work on neural representation of value concerns single cell recordings of midbrain dopamine neurons. These show phasic dopamine increases

with unexpected reward and decreases with unexpected omission of reward (Bayer & Glimcher, 2005; Bayer, Lau, & Glimcher, 2007; Hollerman & Schultz, 1998; Schultz et al., 1997). They thus appear to code for the difference between expected and obtained value., i.e. for an RPE.

It has been claimed by Holroyd and Coles (2002) that it is the phasic activity of midbrain dopaminergic neurons that gives rise to FRN. This component is characterized by a negative-going waveform at medial frontal sites in the interval 200–300 decisions when unfavourable feedback is received. Holroyd and Coles argue that the FRN reflects the transmission of an RPE from the midbrain dopamine system to the ACC. As such, the amplitude of the FRN should be determined by the size of the RPE. In keeping with the theory, subsequent research has shown that unexpected events produce greater FRNs (Bellebaum & Daum, 2008; Hajcak, Moser, Holroyd, & Simons, 2007; Holroyd & Coles, 2002; Potts, Martin, Burton, & Montague, 2006; San Martin, Manes, Hurtado, Isla, & Ibanez, 2010), as do outcomes of greater than expected magnitude (Bellebaum, Polezzi, & Daum, 2010; Kreussel et al., 2012; Marco-Pallares, Kramer, Strehl, Schroder, & Munte, 2010; Wu & Zhou, 2009), though see Hajcak, Moser, Holroyd, and Simons (2006); Yeung and Sanfey (2004).

Both the reinforcement learning theory of the FRN and prospect theory describe valuation. An important difference is that prospect theory describes valuation of outcomes *prior* to their resolution while the FRN purportedly represents valuation *after*. However, insofar as the FRN is believed to represent an RPE, it must formally incorporate (and therefore be sensitive to) the prior valuation term. This is because RPEs consist in the difference between obtained value and EV, and EV is given by the prior valuation of a prospective outcome (in conjunction with its probability). This means the FRN amplitude can be used to retrospectively infer the prior valuation placed on prospects. For example, if you flip a coin to either lose a pound or lose nothing and

you wind up losing the pound, this outcome constitutes an RPE of -£0.50. However, if your prior valuation of the loss of a pound was in fact -£2 (as claimed by prospect theory) then the outcome constitutes an RPE of -£1. Consequently, where FRNs differ despite no difference in the objective value of RPEs we can infer distorted prior valuation.

While the existence in the brain of a biased estimate of EV is implied by the ubiquity of anchor point effects on behaviour, it does not follow that this is used in the generation of RPEs. The discussion of fMRI findings above suggests the possibility of multiple representations of EV, and while anchor point effects are certainly present at the generation of behaviour, it is unclear at what point they are introduced neurally. It might be relatively late, at the point of deciding between alternative actions.

Alternatively, anchor point biases might be applied at the very outset of EV computation, prior to any involvement of systems concerned with behavioural output. Since, as noted above, the FRN should be sensitive to any anchor point effects that are present in the EV term it uses to construct RPEs, it can be used to investigate when, in the route from valuation to behaviour, these anchor point effects are applied. Because the FRN concerns evaluation of feedback, rather than choice between actions, if it reveals anchor point effects then this suggests that these are widespread in the brain's representation of EV. If they are absent it might suggest that they are introduced later and closer to the point of action selection, or reflect the operation of a quite separate valuation system.

In the following experiments, participants undertook a number of gambles which could lead to a good or a bad outcome with respect to the gamble's EV. The monetary difference between the good and the bad outcome was held constant throughout each experiment. This was important because it led to the null hypothesis that differences in the feedback ERP following good and bad outcomes would remain constant regardless

of other experimental factors. This null hypothesis is based on the assumption that the FRN uses an EV term that is unaffected by the zero anchor point and thus codes obtained value and expected value objectively.

In Experiment 1 we manipulated the factor of *domain*, i.e. whether the sum gambled for was gained or lost from participants' bankrolls. Gambles either occurred in the gain domain so that participants won 10 units or won nothing, or they occurred in the loss domain so that participants lost 10 units or lost nothing. The phenomenon of loss aversion is thus a domain effect. While gambles in the gain domain are clearly preferable to those in the loss domain, the *difference* between the good and bad outcomes is the same in both gambles, objectively speaking. Loss aversion, in contrast, predicts that the perceived difference will be greater in the loss domain. If this applies to the FRN then a greater difference should be seen between the waveforms for good and bad outcomes in the loss (-10/0) domain gamble than the win (0/10) domain gamble.

In Experiment 2 we manipulated the factor of *zero proximity*, i.e. whether the sum gambled for lay close to or far from zero. Gambles were either zero-proximal such that participants won 10 units or nothing or they were zero-distal such that participants won 100 units or 90. The phenomenon of diminishing marginal utility is thus a zero-proximity effect. Gambles that are zero-distal are clearly preferable to those that are zero-proximal, however once again, the difference between the good and bad outcomes is the same in both gambles. Diminishing marginal utility predicts, however, that the perceived difference will be greater for zero-proximal gambles. If this applies to the FRN then a greater difference should be seen between the waveforms for good and bad outcomes in the zero-proximal (0/10) gamble than the zero-distal (90/100) gamble. The generation of prospect theory effects from the underlying value function is shown in

Figure 1.

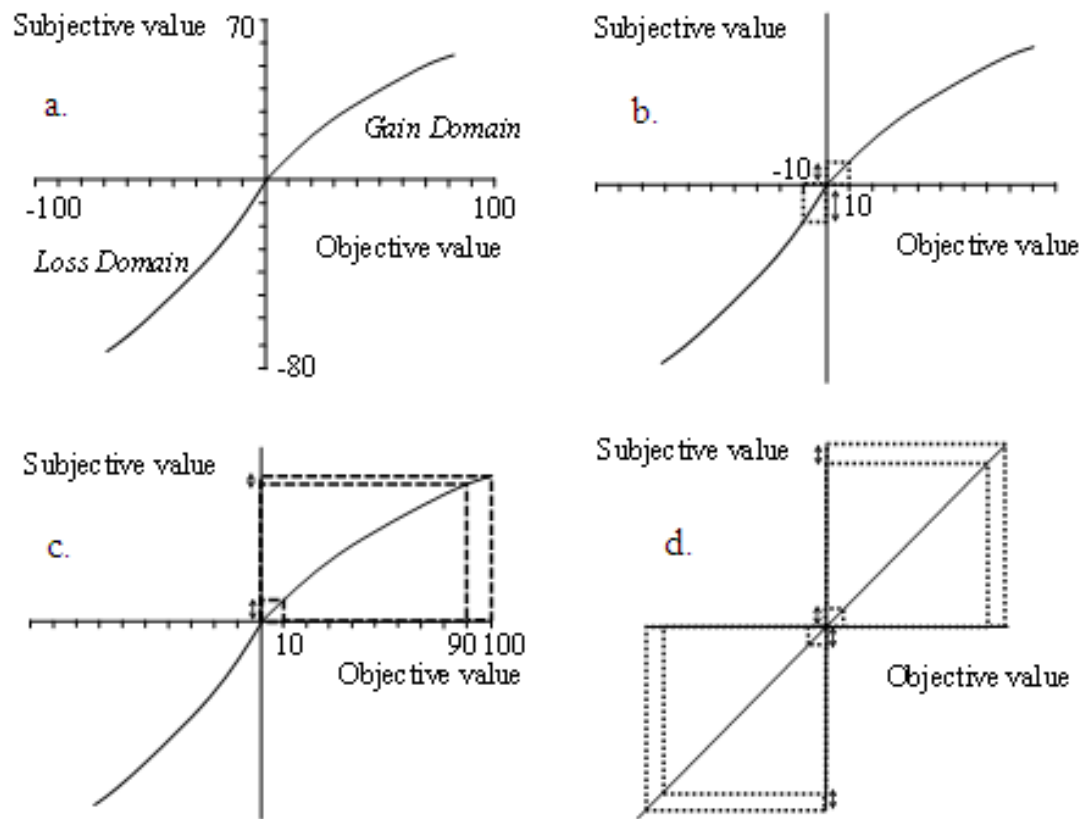


Figure 1 The prospect theory value function

The prospect theory value function a) The prospect theory value function is curved and is steeper for losses. b) A loss of 10 units has more subjective impact than a gain of 10 units as shown by the double-headed arrows. c) A gain of 10 units over an initial nothing has greater subjective impact than a gain of 10 units over an initial 90. d) A value function where there is no zero anchor point: direction and distance from zero have no effect on subjective value.

Regarding the zero-proximity manipulation, we know of no FRN studies that have studied this. Regarding domain, Holroyd, Larsen, and Cohen (2004) found no evidence of this effect when gambles of equal stakes were offered either side of zero. However, looking for loss aversion was not the object of their study and these authors blocked the gambles in the gain and loss domains separately. Given the importance of framing effects in prospect theory (Tversky & Kahneman, 1981), we would expect this to strongly reduce or eliminate the salience of the zero anchor point and attenuate loss aversion. Kreussel et al. (2012) made a comparison of gambles in the loss and gain domains as part of a wider study of the determinants of the FRN. These authors found no effect of domain unless time (first vs. second half of the experiment) was included as a factor, in which case FRN showed increasingly greater sensitivity to gambles in the gain domain. This is an intriguing result, contrary to prospect theory. Nevertheless, the requirement of time as a mediating variable, and the fact that a further two other variables were manipulated in Kreussel et al.'s experiment, increasing the danger of component overlap and also perhaps diminishing the salience of the zero point, suggests further exploration is appropriate. In addition, we also report P3 amplitudes since this component has been shown to be sensitive to the determinants of EV (e.g. Wu & Zhou, 2009) and the stimulus-preceding negativity (SPN), a slow negative wave that precedes stimuli containing information and which is believed to be an index of attention (Brunia, van Boxtel, & Bocker, 2011).

2.3 Experiment 1

2.3.1 Method

Participants. Seventy two undergraduates (3 left-handed, 10 male) participated for course credit and the opportunity to win a small sum of money. Data from six

participants were rejected (three for equipment failure, two for artefacts on over 75% of trials, one for blinking selectively in response to bad feedback). Note: here and in all other experiments, N was set as the maximum sample size possible under the participation point system at that point in the academic year“

Task. Participants were provided with a starting bankroll. To encourage a sense of ownership of this money, participants “won” it in a reaction time task unrelated to the main experiment. While winnings were ostensibly related to reaction times, all participants in fact won £2.90 .

The primary task was presented on a computer using E-Prime software and is depicted in Figure 2. Participants were shown a graphic depicting four squares each containing a symbol. Two of the squares were highlighted with a thick border. Participants were told that on each occasion one of the highlighted symbols would be good and the other bad. They selected one of the two available symbols using keys corresponding to the icon’s position on the screen.

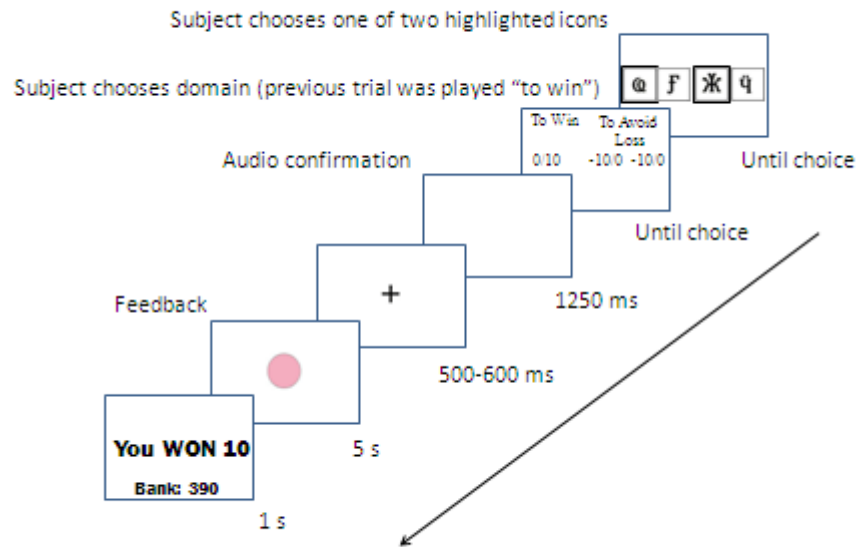


Figure 2 Sequence of stimuli in Experiment 1

Sequence of stimuli in Experiment 1. Participants chose a highlighted icon, selected a domain to play in, received audio confirmation of their choice, fixated, received color-coded feedback on whether a good or bad outcome had been achieved and were shown a summary of winnings so far.

Prior to learning whether their choice was correct, participants were instructed to choose which domain to play in, i.e. “for a win” or “to avoid loss”. This was accomplished with a graphic showing these phrases either side of the screen. On the first of any set of four trials these phrases appeared in duplicate. Selecting an option resulted in one fewer of the corresponding phrase appearing on the next trial. When both copies of a phrase had disappeared, choosing those stakes was not permissible. This ensured that in a block of four trials participants played twice for a win and twice to avoid loss in an order of their choosing. Participants were made to actively choose the domain to increase the perceived relevance of this key manipulation and encourage its retention up to the point of feedback. For the same reason participants chose it *after* they selected the symbol that ostensibly determined whether the outcome would be good or bad. Note that while the order in which participants chose gain vs. loss trials was potentially informative regarding their affective appraisal of gains and losses, it was felt this begged too many questions to permit reliable conclusions: these data were therefore not analysed. An auditory confirmation of participants’ choice of domain lasted 1250 ms, immediately followed by a fixation cross presented for 500–600 ms. Feedback was then presented for 5 s in the form of a pink or yellow circle, denoting whether the good or bad symbol had been chosen (the color representing the good symbol was counterbalanced across participants). Finally, a summary window appeared for 1 s detailing that trial’s winnings and money earned so far. “To win trials” resulted in a win of £0.10 or nothing, “to avoid loss” trials resulted in a loss of £0.10 or nothing. Outcomes were pseudorandomly determined to be good or bad half the time for each domain. The position of the four symbols, and which were highlighted, was randomly determined on each trial. While only two symbols were necessary to implement the experiment, extra symbols were used to obscure the random nature of the task and to

encourage the formation of strategies that would lead to greater interest in feedback (Yeung et al., 2005).

To check that participants were paying attention, both to the domain played in, and the valence of the feedback, after approximately ten percent of trials (randomly determined) they were prompted to press a key to indicate what domain the preceding trial had been played in and what its outcome had been. A 2 s time limit was placed on each response. Participants' bankrolls were increased or decreased £0.10 for right or wrong answers. Total earnings for the experiment averaged £4.62 per person.

Participants played through 192 trials in four blocks separated by 30s breaks. At the experiment's conclusion the satisfaction of the participant with the various outcomes was measured. Participants were shown a graphic depicting a 25 cm line and told one end represented "very unhappy" and the other "very happy". They were asked to position markers on the line to indicate how, on average, the good and bad outcomes for each of the two domains had made them feel. These were converted to numerical scores. The order in which these four judgments were made was counterbalanced across participants.

EEG recording and analysis. EEG data were collected from 11 actively amplified Ag/AgCl electrodes (actiCAP, Brain Products, Gilching, Germany) mounted on an elastic cap. The electrodes were Fz, FCz, Cz, CPz, Pz, F3, F4, FC3, FC4, FP1, FP2. Since the primary focus of this study was the FRN, which is reliably maximal at medial sites, the lateral sites were included merely to allow a preliminary qualitative check that the component identified as the FRN was not greater laterally than medially. Electrodes were referenced to the left mastoid and re-referenced off-line to the average of left and right mastoid activity. Vertical eye movement was monitored by electrodes FP1 and FP2 and a right suborbital electrode, and horizontal eye movement was monitored using an electrode on the right external canthus. Electrode impedances were

kept below 20 k Ω . EEGs were amplified using a BrainAmp amplifier (Brain Products), continuously sampled at 500 Hz, and filtered offline with a band-pass filter from 0.1 to 30 Hz. ERPs were computed by averaging artifact-free EEGs (>90% trials). EEGs were rejected if FP1, FP2 or ocular electrodes showed a voltage change exceeding 100 μ v/100 ms or if any midline site showed a voltage change exceeding 20 μ v/ms.

For the FRN and P3, EEGs were time-locked to 200 ms before the onset of the feedback to 700 ms afterward, and then were baseline-corrected using the period -100 to 0 ms. A problem that has hampered FRN research is lack of consistency in how the component is measured, which can result in the same waveform producing markedly different measures of FRN amplitude depending on the measure. Each method has its relative merits, as discussed by Wu and Zhou (2009). To increase the robustness of our conclusions and to facilitate comparison with the existing and future literature we quantified the FRN using three representative methods. (1) A peak to average peak measure was found by locating three peaks of alternating polarity, a negative peak from 150–280 ms, a following negativity from 200–350 ms and a following positivity from 280–500 ms. The FRN was calculated as the difference of the negative peak and the average of the adjacent positive peaks. A peak to average peak measure was preferred to a peak to peak measure that used only the preceding positivity because grand average waveforms revealed that bad outcomes were associated with a late positivity at ~425 ms that was not attributable to the P3, which showed positivity for good outcomes. The same peak to average peak measure of the FRN was used by Gu et al. (2011), Oliveira, McDonald, and Goodman (2007) and Yeung and Sanfey (2004), all of whom present waveforms in which a late positive deflection is similarly associated with bad outcomes. (2) A mean voltage measure was taken in the interval 240–340 ms, since the meta-analysis in Chapter 5 suggests this is a representative interval. For this measure, the data were high-pass filtered at 2 Hz in order to remove activity associated with the P3 (Wu

& Zhou, 2009). (3) A difference wave measure was taken by finding the peak of the difference wave of good and bad outcomes in the interval 200–300 ms.

The P3 was measured as the peak positivity in the interval 275–500 ms (Yeung & Sanfey, 2004). For the SPN, EEGs were time-locked to 1250 ms prior to the onset of the feedback (500–600 ms following selection of domain) and baseline-corrected using the period -1250 to -750 ms. The SPN was measured as the mean voltage in the interval -200–0 ms preceding feedback onset (Kotani et al., 2003). When analyzing the data, Greenhouse-Geisser corrections were applied where appropriate.

2.3.2 Results

Behavioural results. Performance on the randomly delivered probe questions showed that participants' attention was high (errors recalling domain < 5%, outcome < 4%).

The affective ratings that participants gave at the end of the experiment for the four outcomes are shown in Figure 3. A 2 x 2 ANOVA using the factors of valence and domain revealed greater happiness for good outcomes, $F(1,65) = 441.67$, $p < .001$, $\eta^2 = .87$, for outcomes in the gain domain, $F(1,65) = 89.78$, $p < .001$, $\eta^2 = .58$, and an interaction between these factors $F(1,65) = 9.48$, $p < .01$, that revealed the valence effect to be stronger in the loss domain, as predicted by prospect theory. Further pairwise t -tests revealed that there were significant differences ($p < 0.05$) between each of the individual components of this interaction

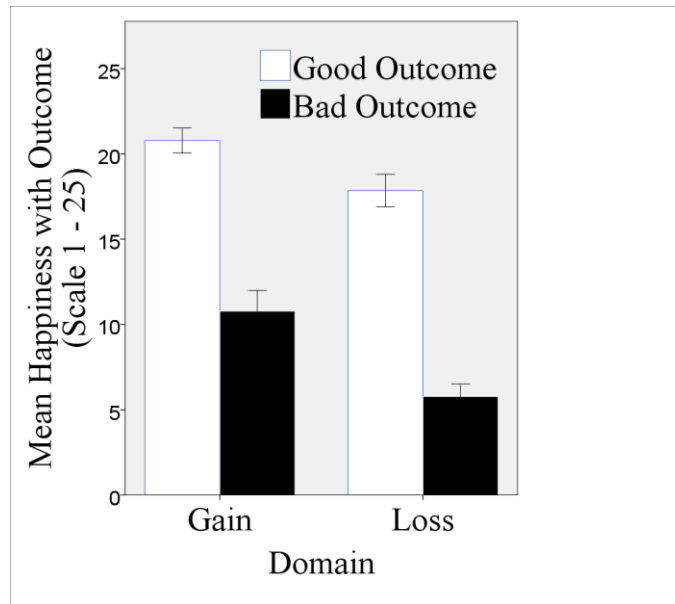


Figure 3 Participants' affective ratings in Experiment 1

Participants' happiness with each of the four possible domain x valence outcomes as measured by self-report at the end of the experiment.

Electrophysiological results. Feedback-related ERP activity typically consists in a frontocentral FRN between 200 and 350 ms and a following parietal P3, both of which show a relative positivity for good outcomes. The bad – good outcome difference term should thus be negative for both components and this can be seen in both the ERP waveform and topographical maps in Figure 4.

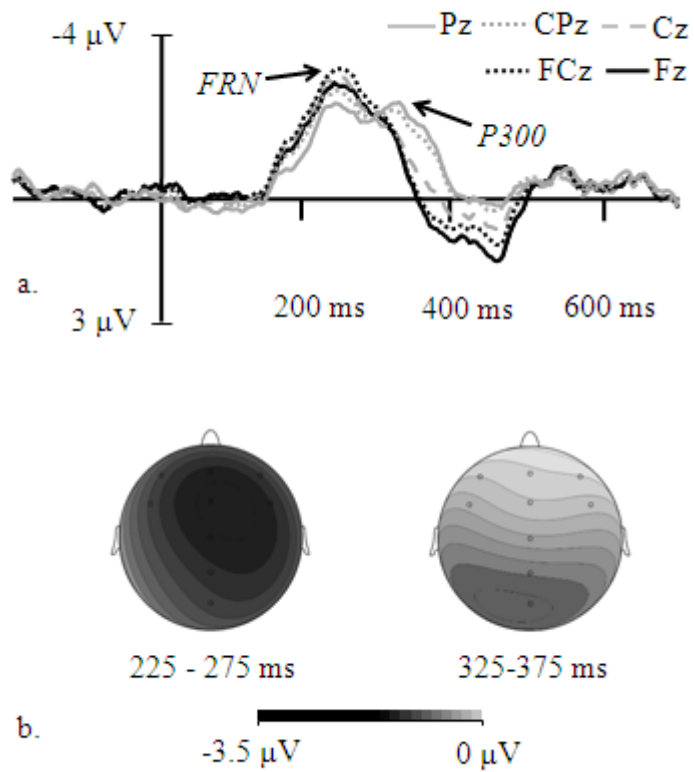


Figure 4 The effect of feedback valence in Experiment 1

The effect of feedback valence (bad – good) in Experiment 1 shown as an ERP waveform (4a) and a scalp topography (4b)

A one way ANOVA testing the effect of midline electrode site (Fz to Pz) on difference wave peak negativity in the interval 200–300 ms revealed a significant effect of site, $F(4,260) = 8.67$, $p < .001$, $\eta^2 = .12$, with the difference wave peak greatest at FCz (4.16 μv) and smallest at Pz (3.97 μv). Consequently, FCz was chosen as the site at which the FRN was measured.

Figure 5 shows the waveforms elicited by good and bad outcomes in the two domains. As described earlier, three FRN measures were employed. A peak to average peak analysis revealed a significant effect of outcome valence, with a greater negative deflection for bad outcomes, $F(1,65) = 4.17$, $p < .05$, $\eta^2 = .06$, and an interaction between valence and domain, $F(1,65) = 8.19$, $p < .01$, $\eta^2 = .12$. Simple effects analysis showed that the valence effect was present in the gain domain, $t(65) = 3.14$, $p < .01$, but not the loss domain, $t(65) = .24$, $p > .05$. Further pairwise t-tests of this interaction revealed that the only significant differences ($p < 0.05$) were as a result of the FRN being significantly higher in the gain/bad condition than all other conditions. There was no main effect of domain, $F(1,65) = 2.05$, $p > .05$.

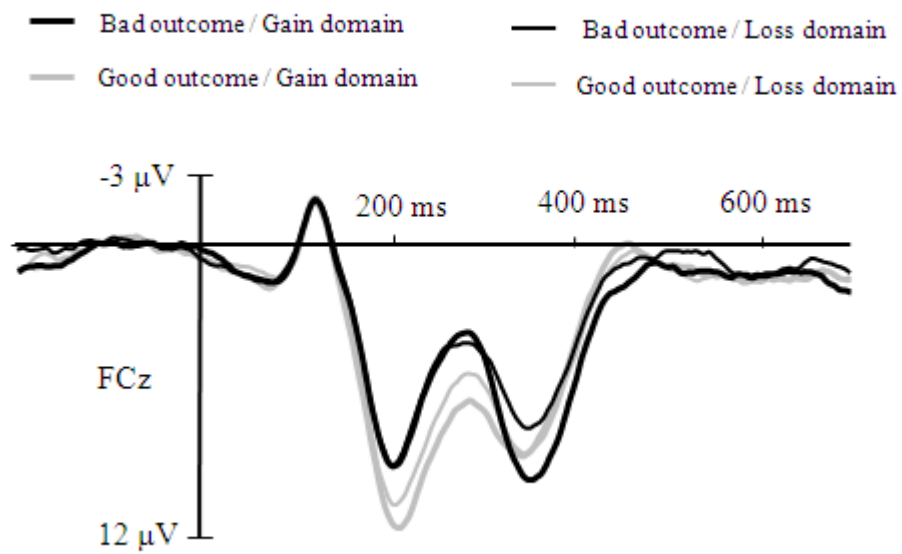


Figure 5 FRN in Experiment 1

FRN in Experiment 1. The figure shows grand average ERPs following feedback on whether a good or bad outcome had been achieved from a monetary gamble. Good outcomes are shown by light lines, bad outcomes by dark lines. Gambles in the gain domain are shown by thick lines, gambles in the loss domain by thin lines. Feedback is at 0 ms. Electrode site is FCz.

A mean voltage measure of FRN revealed a similar pattern of results: significantly more negative voltage for bad outcomes, $F(1,65) = 23.78$, $p < .001$, $\eta^2 = .27$, and a significant interaction between valence and domain, $F(1,65) = 6.90$, $p = .001$, $\eta^2 = .10$. However, simple effects analysis showed that the valence effect was present for both gain, $t(65) = 5.13$, $p < .001$, and loss, $t(65) = 3.88$, $p < .001$ domains. There was no main effect of domain, $F(1,65) = 2.40$, $p > .05$. Further pairwise t-tests revealed that there were significant differences ($p < .05$) between each of the individual components of this interaction with the exception of the gain/bad and loss/bad comparison.

Figure 6 shows bad – good outcome difference waves for the two domains. The difference wave peak in the gain domain was significantly greater, $t(65) = 3.12$, $p < .01$, thus supporting the interaction between domain and valence shown in the previous measures.

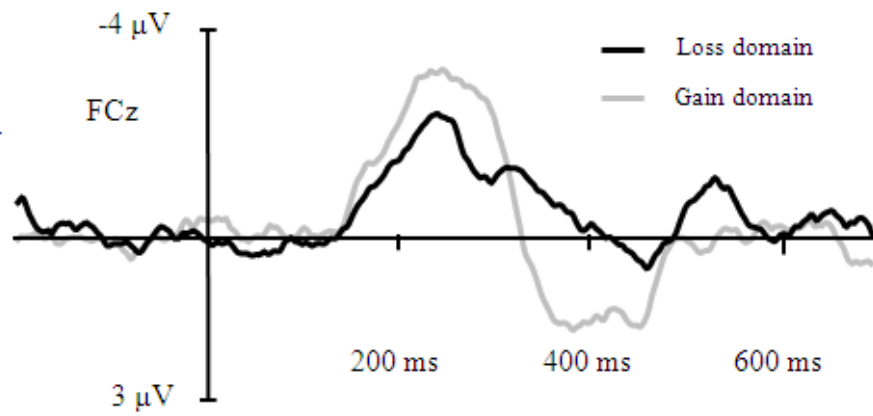


Figure 6 Difference waves in Experiment 1

Difference waves created from waveforms shown in Figure 5. The good outcome waveform is subtracted from the bad outcome waveform for each domain separately.

Following Kreussel et al. (2012), we examined the effect of time by splitting the data into two halves. Three participants were removed because trials fell below the recommended minimum of 20 (Marco-Pallares, Cucurell, Munte, Strien, & Rodriguez-Fornells, 2011). Using a peak to average peak analysis, the three way domain x valence x time interaction fell just short of significance $F(1,62) = 3.36$, $p < .10$. We explored this further by checking whether the valence x domain interaction shown previously was present in both halves of the experiment. There was no interaction in the first half, $F(1,62) = 1.26$, $p > .05$, but a strong interaction in the second, $F(1,62) = 12.53$, $p < .001$, $\eta^2 = .17$. Using a mean voltage measure, no significant time x valence x domain interaction was found ($F < 1$).

Regarding the P3, a one way ANOVA testing the effect of midline electrode site (Fz to Pz) on the positivity of the waveform (all conditions averaged) in the interval 275–500 ms revealed a significant effect of site, $F(4,260) = 8.67$, $p < .001$, $\eta^2 = .13$, with the peak positivity greatest Pz (10.52 μv) and smallest at Fz (7.40 μv). Consequently, Pz was chosen as the site at which the P3 was measured. The P3 is shown in Figure 7. An analysis of the P3 revealed that its amplitude was significantly greater for good outcomes, $F(1,65) = 25.94$, $p < .001$, $\eta^2 = .29$, and for outcomes in the gain domain, $F(1,65) = 8.93$, $p < .01$, $\eta^2 = .13$, but there was no interaction between valence and domain, $F < .01$.

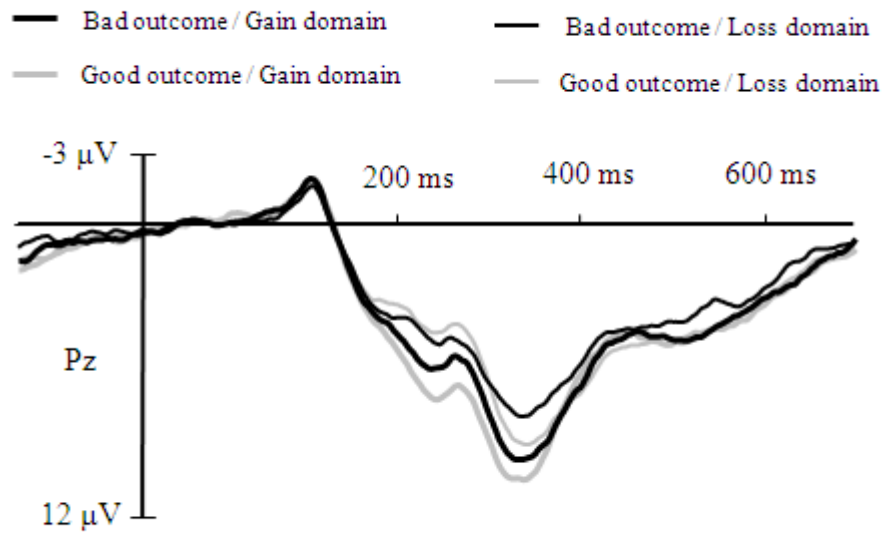


Figure 7 P3 in Experiment 1

P3 in Experiment 1. The figure shows grand average ERPs following feedback on whether a good or bad outcome had been achieved from a monetary gamble. Good outcomes are shown by light lines, bad outcomes by dark. Gambles in the gain domain are shown by thick lines, gambles in the loss domain by thin lines. Feedback is at 0 ms. Electrode site is Pz.

Regarding the SPN, a one way ANOVA testing the effect of midline electrode site (Fz to Pz) on the mean negativity of the waveform (all conditions averaged) in the interval -200 to 0 ms pre-feedback revealed a significant effect of site, $F(4,260) = 16.34$, $p < .001$, with greatest negativity at Cz ($-7.48 \mu\text{v}$) and least at Fz ($-5.28 \mu\text{v}$). Consequently, Cz was chosen as the site at which the SPN was measured. Results revealed that the SPN was significantly greater preceding loss domain feedback, $t(65) = 2.46$, $p < .05$.

2.3.3 Discussion

The stated satisfaction of participants with outcomes showed strong loss aversion in that the difference between good and bad outcomes was perceived to be greater when these occurred in the loss domain. In keeping with this, the SPN, an index of anticipatory attention, was greater for loss domain gambles, suggesting the forthcoming outcome was perceived as more important. However, in the FRN, the reverse effect was found: increased sensitivity to valence effects in the gain domain. The P3 was elevated for outcomes that were good, and for those in the gain domain but these factors did not interact.

2.4 Experiment 2

2.4.1 Method

Participants. Fifty-nine undergraduates (5 left-handed, 24 males) participated for course credit and the opportunity to win a small sum of money. No participants were subsequently rejected.

Task. The task was essentially similar in structure to Experiment 1 with the following differences. As domain and marginal utility terms are orthogonal in Kahneman and Tversky's (1979) mathematical realisation of prospect theory, the gambles in this experiment were presented in a single domain. The gain domain was chosen, as the effect of valence was found to be stronger in the gain than the loss domain in Experiment 1, whilst also being the most commonly studied domain within the literature. Since participants could not incur negative winnings, the reaction time task from Experiment 1 was omitted. In the main experimental task, in order to provide winnings that were either close or far from the zero anchor point, and at the same time keep participant payments manageable, the game was played for points. On a given trial, participants could either earn 0 or 10 points, or they could earn 90 or 100 points. Participants were advised that these would be later converted to cash at the rate of £0.01 per 25 points. All other events on a trial resembled Experiment 1 except that instead of choosing to play "to win" or "to avoid loss", participants chose gambles that were either close to or far from zero, namely "nothing or ten" or "ninety or a hundred". These choices were displayed as "0/10" and "90/100". Total earnings for the experiment averaged £4.82 per person.

2.4.2 Results

Behavioural results. Performance on the randomly delivered probe questions showed that participants' attention was high (errors recalling zero-proximity < 5%, outcome < 3%). The affective ratings that participants gave at the end of the experiment for the four outcomes are shown in Figure 8. A 2 x 2 ANOVA of zero-proximity and outcome valence was performed. This revealed greater happiness for good outcomes, $F(1,58) = 341.52, p < .001, \eta^2 = .78$ for zero-distal outcomes (which was unsurprising, since these were larger sums) $F(1,58) = 131.85, p < .001, \eta^2 = .61$ and an interaction, $F(1,58) = 13.21, p < .001, \eta^2 = .16$, such that the valence effect was stronger for the

zero-proximal outcomes than the zero-distal ones, as predicted by prospect theory.

Further examination using pairwise t-tests revealed that there were significant differences ($p < 0.05$) between each of the individual components of this interaction with the exception of the zero-proximal/good vs. the zero-distal/bad outcome.

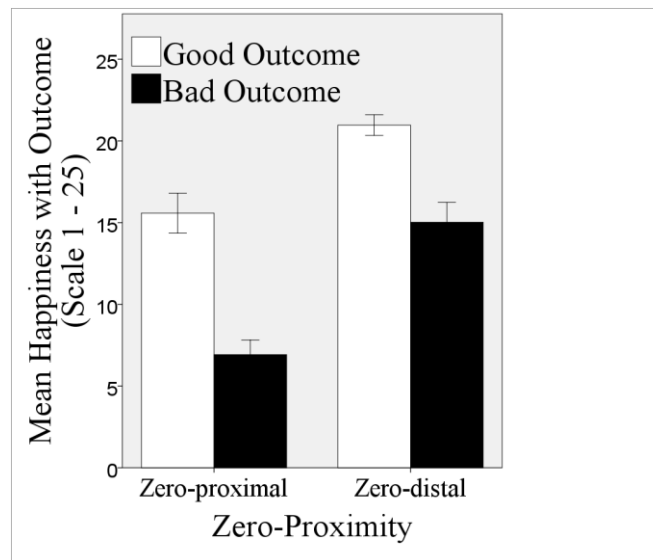


Figure 8 Participants' affective ratings in Experiment 2

Participants' happiness with each of the four possible zero-proximity x valence outcomes as measured by self-report at the end of the experiment.

Electrophysiological results. Figure 9 shows bad - good outcome voltage differences. It shows the same progression of FRN to P3 seen in Experiment 1. The site at which FRN was maximal (using the negative peak of the difference wave at 200–300 ms) was once again established using a one way ANOVA using the five midline electrodes. Whilst this did not produce a significant effect of site, ($F < 1$), the difference wave peak was greatest at FCz (4.14 μv) and smallest at Pz (3.96), and thus FCz was chosen as the site at which the FRN was measured.

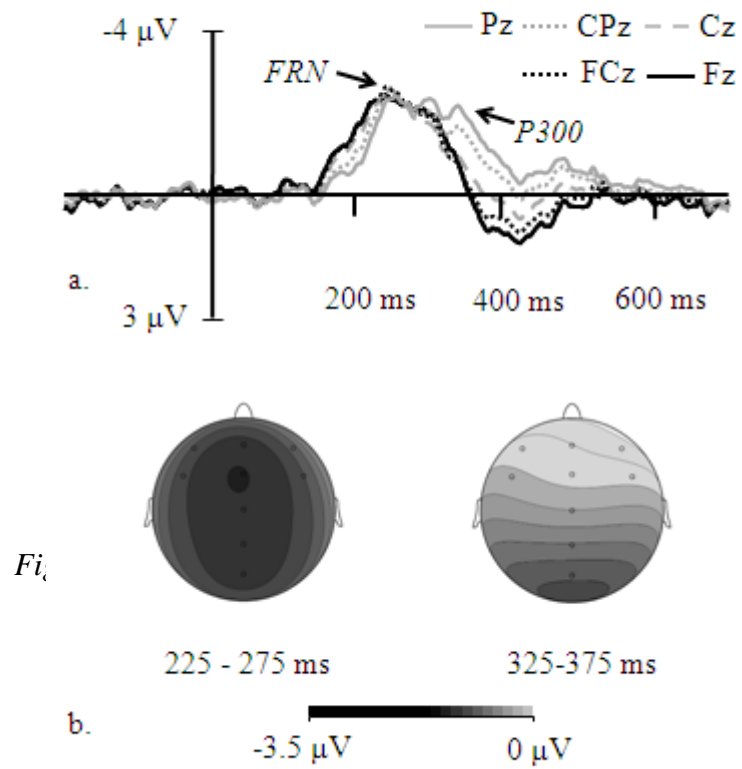


Figure 9 The effect of feedback valence (bad – good) in Experiment 2 shown as an ERP waveform (9a) and a scalp topography (9b)

Figure 10 shows the ERPs elicited by good and bad outcomes for zero-proximal (0/10) and zero-distal (90/100) gambles. A peak to average peak analysis revealed a significant effect of valence, with greater negative deflection for bad outcomes, $F(1,58) = 9.21, p < .01, \eta^2 = .12$. There was no effect of zero-proximity, $F < .1$, nor was there a significant interaction between zero-proximity and valence, $F < .1$. A mean voltage measure of FRN revealed the same pattern of results: a significantly more negative waveform for bad than good outcomes, $F(1,58) = 24.25, p < .001, \eta^2 = .30$, but no main effect of zero-proximity, $F(1,58) = 2.41, p > .05$, or interaction with valence, $F < .1$. Figure 11 shows bad - good outcome difference waves, with the peak of difference wave measure of FRN in agreement with the previous analyses in showing no effect of zero-proximity, $t < 1$.

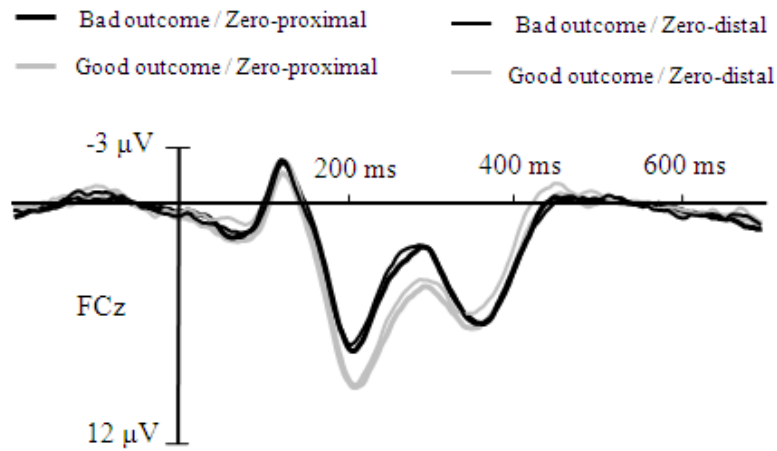


Figure 10 FRN in Experiment 2

FRN in Experiment 2. The figure shows grand average ERPs following feedback on whether a good or bad outcome had been achieved from a monetary gamble. Good outcomes are shown by light lines, bad outcomes by dark lines. Zero-proximal gambles are shown by thick lines, zero-distal gambles by thin lines. Feedback is at 0 ms. Electrode site is FCz.

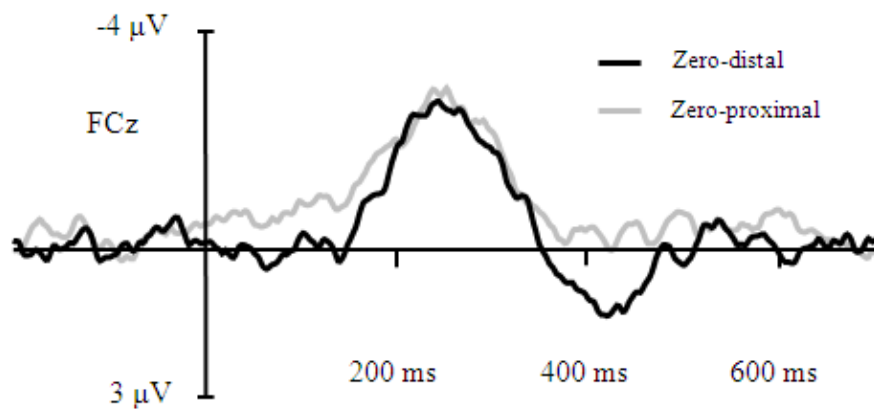


Figure 11 Difference waves in Experiment 2

Difference waves created from waveforms shown in Figure 10. The good outcome waveform is subtracted from the bad outcome waveform for the zero-proximal and zero-distal gambles separately.

Regarding the P3, a one way ANOVA testing the effect of midline electrode site (Fz to Pz) on the positivity of the waveform (all conditions averaged) in the interval 275–500 ms revealed a significant effect of site, $F(4,232) = 34.49$, $p < .001$, $\eta^2 = .33$, with the peak positivity greatest at Pz (8.72 μv) and smallest at Fz (5.32 μv). Consequently, Pz was chosen as the site at which the P3 was measured. An analysis of the P3, shown in Figure 12, revealed that the amplitude of this component was significantly greater for good outcomes, $F(1,58) = 18.12$, $p < .001$, $\eta^2 = .21$, but that there was no effect of zero-proximity, $F < 1$, or interaction with valence, $F < 1$.

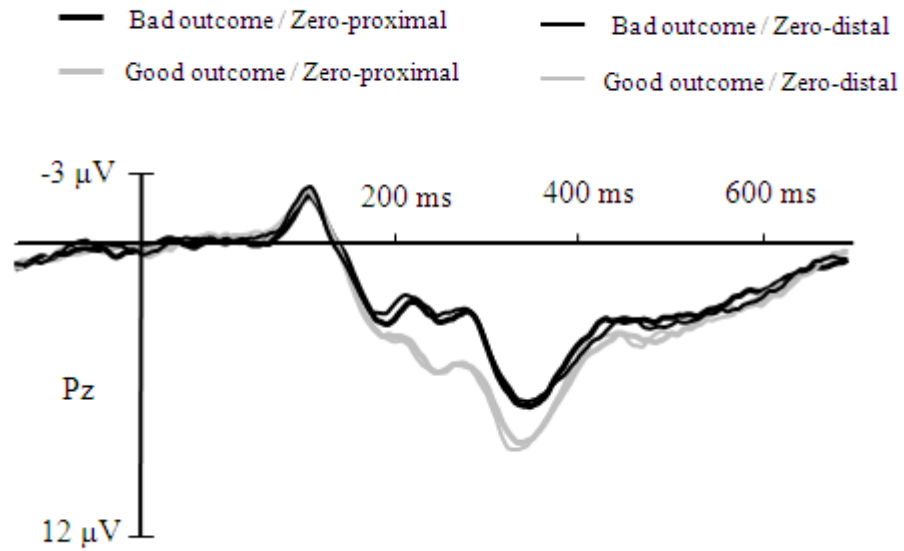


Figure 12 P3 in Experiment 2. The P3 component of the ERP elicited by monetary feedback on whether a good or bad outcome had been achieved from a monetary gamble. Good outcomes are shown by light lines, bad outcomes by dark. Zero-proximal gambles are shown by thick lines, zero-distal gambles by thin lines. Feedback is at 0 ms. Electrode site is Pz.

Regarding the SPN, a one way ANOVA testing the effect of midline electrode site (Fz to Pz) on the mean negativity of the waveform (all conditions averaged) in the interval -200 to 0 ms pre-feedback revealed a significant effect of site, $F(4,232) = 10.09$, $p < .001$, $\eta^2 = .13$, with greatest negativity at Cz (-6.03 μv) and least at Pz (-4.74 μv). Consequently, Cz was chosen as the site at which the SPN was measured. The SPN showed no significant difference between zero-proximal and zero-distal gambles, $t < 1$.

2.4.3 Discussion

As with Experiment 1, participants' stated satisfaction with outcomes followed prospect theory, showing greater sensitivity to zero-proximal outcomes. However, no such effect was observed in the FRN, which was merely responsive to outcome valence. The P3 was likewise sensitive only to valence. There was no effect of zero-proximity on the SPN.

2.5 General Discussion

People typically succumb to two anchor point dependent biases in value coding, sensitivity to the domain (gain or loss) and sensitivity to zero-proximity (zero-proximal or zero-distal). When people make judgements of the value of prospects, that is, of their expected value, they appear to make them with respect to these biased values.

The rationale of the experiments presented here was to see if these biases of expected value were used in the generation of the FRN. The FRN is believed to represent an RPE, that is, a comparison of obtained value with EV. If the EV term used in the generation of the FRN showed anchor point effects, this would suggest that these were an integral feature of valuation in the brain, not a bias introduced only at the point at which behaviour was implemented.

The biases, evident in participants' affective ratings, were not present in the FRN, suggesting anchor point effects may not be integral to EV representation. There was no effect of marginal utility on the FRN. Our confidence in this null result is bolstered by the large sample size and the fact that essentially the same procedure was capable of eliciting an effect ($\eta^2 = .10$) showing greater sensitivity to gains in Experiment 1. While Experiment 2 had very slightly fewer participants, the close design similarities suggest a similar power value, but in this case the interaction between valence and zero proximity produced a η^2 value of .002 and so we are reasonably confident that these factors do not interact. It should be acknowledged however that certain practical limitations to the experiment leave other interpretations possible. The small value of the sums involved may have meant that participants perceived themselves to be in the linear portion of the value function, removing differential marginal utility effects. It is also possible that scaling effects may have played a role, even though Chapter 5 suggests the evidence for these is weak

In Experiment 1, in contrast to loss aversion, participants' FRNs showed greater sensitivity to gains. There is, to our knowledge, no single cell study of domain effect to which we can compare this result. The FRN is believed to arise when midbrain RPEs reach either the ACC or the basal ganglia, but it would be unwise to assume that this observed gain sensitivity is necessarily present in midbrain coding, since it may arise from modulation of an initially unbiased midbrain RPE signal. In either event however, this study suggests a value function for the FRN that is quite different from the one proposed by prospect theory, that is shown in Figure 1a. While a graded, parametric manipulation of the factors of domain and zero-proximity is ultimately required to map the FRN value function, evidence from this study provisionally suggests a linear rather than curved function (Experiment 2), which is steeper for gambles in the gain domain than in the loss domain (Experiment 1).

An alternative explanation for the observed domain x valence interaction in Experiment 1 might arise from the fact that the gain/good outcome is dissimilar to the other three outcomes in that there is no aspect of that outcome which is negative. It might thus constitute a qualitatively different outcome which drives the interaction and supposed greater sensitivity to valence in the gain domain. However, t-tests revealed that the gain/good outcome was not an outlier. Under a mean voltage measure of the FRN both the gain/good and gain/bad outcomes showed significant differences from other conditions, while under a peak to average peak measure it was the gain/bad condition that was an outlier compared to the other conditions.

A further potential limitation of the experiment is that there may have been greater sensitivity to gains than to losses because participants did not in fact lose their own money in the experiment.

In order to integrate our findings with the existing fMRI literature on prospect theory effects it is necessary to consider the neural generator of the FRN. The ACC has traditionally been believed to be the source of the FRN. However, recent evidence has suggested that at least some portion of the FRN may arise in the striatum. Using both fMRI and ERPs, Carlson, Foti, Mujica-Parodi, Harmon-Jones, and Hajcak (2011) found that FRN correlated with BOLD activation in the ventral striatum, dorsal striatum and medial prefrontal cortex. However, principal components analysis made in that study and others conducted by Foti, Weinberg, Dien, and Hajcak (2011b) and ourselves (Chapter 4) identified the striatum as the most probable generator of the FRN. This was in agreement with the conclusions of Martin, Potts, Burton, and Montague (2009). The issue remains unresolved (Cohen, Cavanagh, & Slagter, 2011; Foti, Weinberg, Dien, & Hajcak, 2011a), however the frequency with which the striatum is found to be activated by reward encoding justifies its consideration as a contributor to the FRN.

Our domain effect stands in direct contrast to the finding of Tom et al.'s (2007) fMRI study. This showed that the striatum responded to outcomes in both gain and loss domains but with increased sensitivity in the loss domain. These authors suggested that the striatum was a source of value coding with an inbuilt loss aversion parameter. Others have proposed a more componential basis for value coding. For example, both Yacubian et al. (2006) and Knutson, Adams, Fong, and Hommer (2001) found the ventral striatum to be responsive only to gains, while losses and loss aversion were associated with activity in the amygdala (De Martino et al., 2006; Weber et al., 2007; Yacubian et al., 2006) or insula (Knutson et al., 2008). A resolution of the present findings and Tom et al.'s concerns the differential roles that might be played by dopamine and serotonin in coding loss and gain domain outcomes. Opponency between dopamine and serotonin is well established (Daw, Kakade, & Dayan, 2002). While the FRN is believed to arise from dopamine, in the striatum there is evidence of a serotonin-dopamine gradient along a caudal-rostral axis (Brown & Molliver, 2000; Heidbreder, Hedou, & Feldon, 1999). Intriguingly, Seymour et al. (2007) found that within the striatum, anterior regions showed relative selectivity for gains, and posterior regions for losses. This accords with a number of recent studies which have suggested that these two neuromodulators contribute differentially to coding for outcomes less than or greater than the zero anchor point. For example, Zhong et al. (2009) have shown that genetic variation in tonic dopamine and serotonin levels modifies risk seeking in gain and loss domains respectively. Pessiglione, Seymour, Flandin, Dolan, and Frith (2006) showed that dopamine agonists affected choices in the gain domain (both neurally and behaviourally) but not the loss domain. Campbell-Meiklejohn et al. (2011) showed that serotonin but not dopamine promoted "chasing behaviour", that is gambles undertaken by participants who consider themselves to currently hold assets below an anchor point (typically what they started a session with).

We also measured the P3, a component believed to be broadly sensitive to the informational content of a stimulus. While it might be expected that the P3 would be greater in the loss domain, as a result of loss aversion, we found that gain domain gambles elicited a larger P3 overall. This accords with Kreussell et al.'s (2012) results (although the P3 fell just short of significance in their study). The P3 is assumed to respond to *task relevant* information (e.g. Nieuwenhuis, Aston-Jones, & Cohen, 2005) and so this component may tell us something about how participants perceived the task they were engaged in. If so, participants appear to have perceived gambles in the gain domain as more relevant to the task at hand than those in the loss domain. Note that while net earnings in Experiment 1 were predetermined to be zero, participants' bankrolls increased over time due to the bonus money received from correctly answering probe questions. Once an upward trajectory began to be accepted as the norm, participants may have come to perceive their task as maximizing the speed with which they increased their bankroll, and while loss domain gambles are logically as relevant as gain domain gambles to this task, participants may not have perceived them as such. Kreussell et al. (2012) suggested that this asymmetry of perceived task-relevance of the two domains could in fact underlie the FRN sensitivity to gains in their study, which was only observable when (growth over) time was added as an extra variable. In the current experiment the support for this time-based interpretation of the domain effect was not strong insofar as the three way time x domain x valence interaction was present for only one of the three methods of quantifying the FRN. Furthermore, evidence against Kreussell et al.'s interpretation comes from the SPN. This component is assumed to reflect anticipatory attention (van Boxtel & Bocker, 2004). In our data the SPN for loss gambles was greater than that for gain gambles suggesting that participants were more sensitive to prospective losses, in keeping with their later affective ratings. This contrasted with their FRNs, where there was greater sensitivity to valence in gain

gambles, as indicated by the greater valence effect in that domain. We therefore do not regard participants' perception of the task as the basis of the domain effect found in the FRN.

In summary, the experiments described in this study revealed a discrepancy between the FRN and conscious affective ratings. Participants' affective ratings were sensitive both to the RPE and the actual outcome value. Sensitivity to outcome value was shown by an overall preference for gain domain gambles and for zero-distal gambles, which was unsurprising in that these gave the better monetary return on average than loss domain gambles and zero-proximal gambles. Sensitivity to RPEs was shown by good outcomes being preferred to bad. These simultaneous effects are nicely demonstrated by participants' ratings of zero-proximal/good and zero-distal/bad conditions in Experiment 2. The zero-distal/bad outcome was bad compared to its alternative yet produced a greater increment in bankroll than the zero/proximal good outcome. The sum effect of these dual considerations on participants' happiness appears to be that these two outcomes were rated equally pleasant. At a conscious, reflective level participants thus appear to rate outcomes using both the outcome's value and the RPE it constituted. These very strong main effects in both experiments were modulated by significant, but weaker interactions consistent with prospect theory. A quite different picture was shown by the FRN, for which the strongest effects arose from valence, and for which there was no effect of zero proximity and a domain effect in the opposite direction. Overall, this suggests a dissociation between valuation performed by the reinforcement learning system that the FRN is presumed to serve and wider, more reflective aspects of cognition. We return to this question in Chapter 5.

3 An Examination Of Sensitivity To Positive, Negative And Unsigned Prediction Errors

The work in this chapter is strongly based on a published paper (Sambrook and Goslin, 2014)

3.1 Chapter Abstract

Reinforcement learning models make use of RPEs, the difference between an expected and obtained reward. There is evidence that the brain computes RPEs, but an outstanding question is whether positive RPEs (“better than expected”) and negative RPEs (“worse than expected”) are represented in a single integrated system. An electrophysiological component, feedback related negativity, has been claimed to encode an RPE but its relative sensitivity to the utility of positive and negative RPEs remains unclear. This study explored the question by varying the utility of positive and negative RPEs in a design that controlled for other closely related properties of feedback and could distinguish utility from salience. It revealed a mediofrontal sensitivity to utility, for positive RPEs at 275–310 ms and for negative RPEs at 310–390 ms. These effects were preceded and succeeded by a response consistent with an unsigned prediction error, or “salience” coding.

3.2 Introduction

The FRN is so named because it exhibits a relative negativity for worse than expected outcomes. However, this does not necessarily mean exclusive, or indeed any, sensitivity to –RPEs. The methodology in which the FRN emerged was based simply on

comparing –RPE outcomes to +RPE outcomes; as such the negativity observed is merely relative and might equally have its basis in a positive voltage shift for +RPE outcomes. In fact, competing claims have been made in this regard, with some arguing that the FRN is preferentially sensitive to +RPEs (Cohen, Elger, & Ranganath, 2007; Eppinger, Mock, & Kray, 2009; San Martin et al., 2010) and others arguing greater sensitivity to –RPEs (Bellebaum & Daum, 2008; Bellebaum, Polezzi, et al., 2010; Pfabigan, Alexopoulos, Bauer, & Sailer, 2011). In Figures 13a and b we schematically represent these two possible response functions. Figure 13c shows the response function of a component that codes both +RPEs and –RPEs (“integrated coding”), and Figure 13d a response function to unsigned prediction errors (UPEs), that is to the absolute size of the prediction error irrespective of its valence. While this last response function has been plotted against RPE utility like the others, this merely represents how it would *behave* in an experiment studying RPEs, the component is not coding RPE utility at all but the quite different property of UPE size. Such a component might serve a general function of registering motivational salience (Bromberg-Martin, Matsumoto, & Hikosaka, 2010).

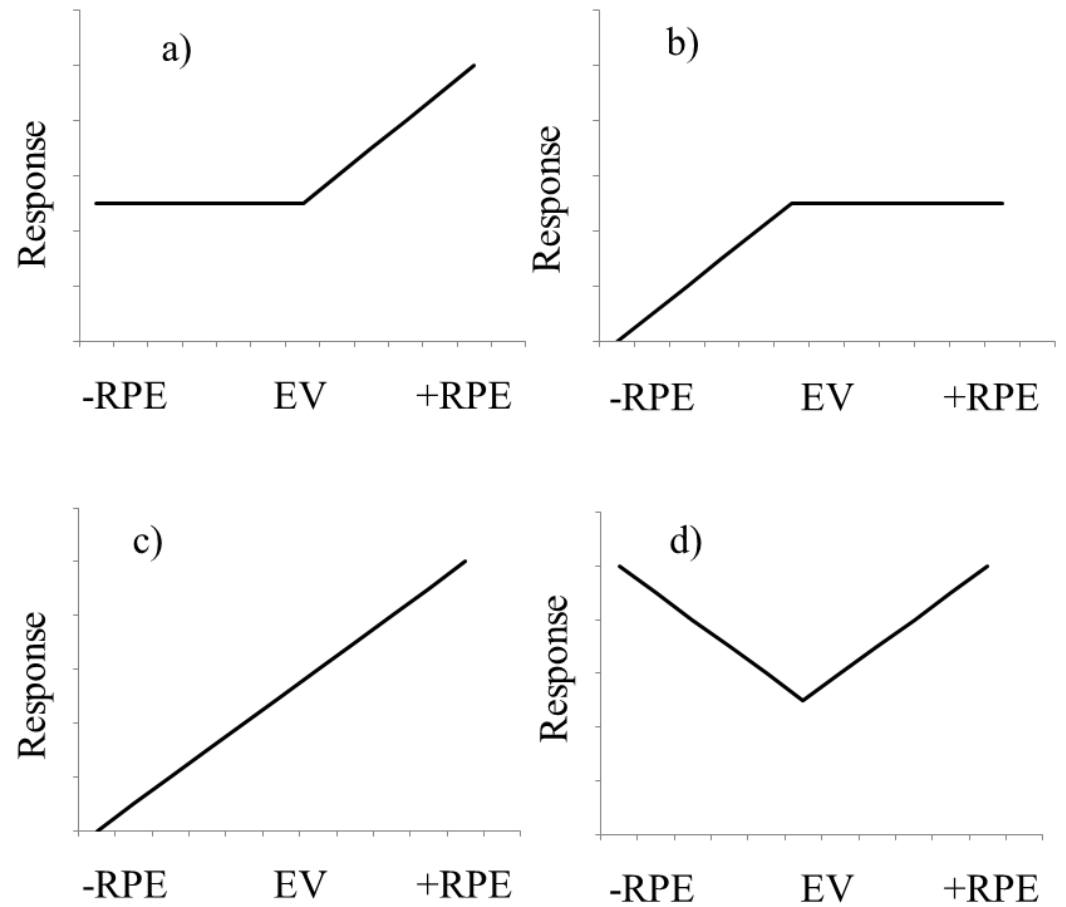


Figure 13 A schematic representation of possible response functions to RPE utility.

A schematic representation of possible response functions to RPE utility. “Response” is generic and could refer to increase in single cell firing rate, BOLD activation, or amplitude increase for either a positive or negative going ERP component. In the specific case of the FRN the y axis corresponds to positivity of voltage. a) coding of +RPE utility only, b) coding of –RPE utility only, c) “integrated” coding of the utility of all RPEs, d) coding of UPE size, or “salience”. EV = expected value, i.e. an RPE utility of zero

A challenge for FRN research is that the post-feedback waveform may comprise a number of different components with different response functions that at least partially overlap. One danger is that this overlap, rather than merely obscuring the individual components, has the capacity to synthesise entirely artefactual response functions. For example if a UPE size response function (Figure 13d) overlies an integrated RPE response function (Figure 13c) the sum effect in the EEG will be a spurious response function corresponding to Figure 13a, a +RPE encoder. Where temporal and spatial overlap of such components is perfect, this problem is insoluble for the ERP methodology. However, even when two components might in principle be temporally dissociable, in practice this distinction may fail to be made since components are generally quantified over a relatively wide interval of 100 ms or so, a problem we address in the present experiment.

The example of a spurious +RPE encoder given above is topical. The recent trend in FRN research, following Holroyd et al. (2008), has been to claim that the FRN is solely responsive to +RPEs, despite the clear loss of adaptive value this would hold relative to an integrated RPE encoder. Tellingly, there has also been recent growth in the number of papers claiming that the FRN does not code RPE utility at all and that is in fact just a UPE size response (Hauser et al., 2014; Oliveira et al., 2007; Talmi et al., 2013). If the post-feedback waveform comprises an early UPE size response closely followed by an integrated RPE utility response then an interval of measurement that catches part of both these responses will generate the spurious +RPE encoding response function as described.

The approach we took in the present study was therefore to quantify components according to the activity present in actual data using a ‘bottom-up’ strategy, rather than

imposing a possibly misguided interval of measurement (on which there is little agreement in the literature). We also depart from the tendency in FRN studies to represent RPEs as categorical levels such as “large good outcomes” vs. “small good outcomes” so that they can fit into a factorial design. This limitation belies the essentially continuous nature of RPEs. In our experiment we manipulated RPEs as a continuous independent variable, analysing the effect of RPE utility on voltage by *correlating* the two variables. Where the correlation between these two values was found to be significant, it could be assumed that voltage was influenced by RPE utility, thus indicating the presence of an RPE encoding component. This relatively novel technique has been used by Hauk, Davis, Ford, Pulvermuller, and Marslen-Wilson (2006). The appropriateness of correlation coefficients for answering our research question can be readily appreciated by a glance at Figure 13. The response functions there are schematic, however with the addition of a scatterplot of data points, each subplot might as easily represent two lines of best fit (one for +RPEs and one for –RPEs) set end to end, with the closeness of scatter on each line corresponding to two correlation coefficients for the responsiveness of voltage to +RPEs and –RPEs respectively. Using such a correlational waveform has a number of advantages: it harnesses the extra power of parametric designs over factorial ones and it tests for a monotonic relationship between a continuous IV and a component’s amplitude, something frequently assumed, but not demonstrated, in factorial designs. It greatly simplifies presentation of results, since instead of plotting ERPs at each level of a factor (30 levels in the present case), a single waveform is plotted.

Practically, identification of RPE encoding components in the post-feedback waveform, and elucidation of their response functions was achieved in two stages, as follows. Separate correlations of voltage with +RPE utility and –RPE utility were calculated at each time point on the post-feedback waveform. Separating the two kinds

of RPE at the correlation stage was essential: if the correlation were calculated over the full range of RPEs then a significant positive value (for example) of Pearson's r would leave us uninformed as to whether we were observing a response function corresponding to Figure 13a, b or c. Having generated these correlational waveforms, the waveforms were clustered into discrete intervals which were tested for significance using the procedure of Maris and Oostenveld (2007). This showed us the intervals where $-RPE$ and $+RPE$ encoding was occurring, but not the overall response functions in these intervals. This was achieved in the second stage by considering $+RPE$ and $-RPE$ encoding together and comparing the joint activity to the response functions in Figure 13. For example an interval of the waveform where there was a strong correlation between $+RPE$ utility and voltage could be described as a $+RPE$ encoder only if there was no correlation between $-RPE$ utility and voltage in that interval. If there was in fact a same-signed correlation between $-RPE$ utility and voltage then this pattern of results suggested an integrated encoder. If there was an oppositely-signed correlation between $-RPE$ utility and voltage in this interval then this suggested a UPE size encoding.

These novel measures were designed to deal with the likely presence of multiple components in the post feedback waveform, and maximise the possibility that they might be separated out into their true response functions. An essential requirement of this endeavour was that the experimental design should avoid confounds seen in the existing literature that may have served to distort the apparent sensitivity of the FRN to $+RPE$ s and $-RPE$ s, thus masking the FRN's true response function. One of these is the domain of the outcome, as the FRN was shown to be sensitive to this factor in Chapter 2. In the present experiment, expectations were manipulated in such a way that $-RPE$ s and $+RPE$ s were both expressed as losses and gains half the time, recovering the orthogonal nature of the relationship of outcome domain and RPE valence.

A second important confound concerns whether outcomes are deliveries or omissions. When FRN experiments do not use mixed gambles, they near-ubiquitously offer gambles in the gain domain only and modulate RPE utility by whether a reward has been obtained or not. Thus a likelihood-modulated FRN experiment might offer outcomes as follows: 25¢ when reward was unlikely, 25¢ when reward was likely, nothing when non-reward was likely, nothing when non-reward was unlikely; alternatively, a magnitude modulated FRN experiment would offer 25¢, 5¢, non-reward (where the alternative was 5¢), non-reward (where the alternative was 25¢). However, it has been argued that non-rewards are less salient than rewards (Miller, Barnet, & Grahame, 1995). If this is so then an apparent lack of sensitivity to the utility of –RPEs may occur simply because they have been expressed as non-rewards. In the present experiment there were no non-rewards: all outcomes constituted deliveries of some numerical quantity, with –RPEs being worse than expected quantities.

A third possible source of confounds surrounds the use of reward likelihood (rather than reward magnitude) to manipulate RPE utility. Although this has historically been the preferred means of manipulating RPE utility in FRN experiments, unexpected events are known to have very strong effects in the time course of the FRN (Folstein & Van Petten, 2008) and while unexpectedness may play a formal role in dictating the utility of RPEs, it is likely to bring with it substantial non-specific, alerting responses. For this reason, the present experiment manipulated RPEs using outcome magnitude rather than likelihood. Using reward magnitudes of equal frequency preserves the formal manipulation of RPE utility, but removes non-specific surprise effects.

A related, though more insidious confound, may exist in the form of perceptual mismatch between an expected, or hoped for stimulus and the actual feedback. Jia et al. (2007) have shown that a negativity is elicited in the FRN interval when a stimulus differs from one predicted by a participant, regardless of whether match or mismatch is

the winning criterion, while Donkers, Nieuwenhuis, and van Boxtel (2005) showed that when a stimulus breaks a pattern with an ongoing sequence this leads to a net negativity even when this mismatch denotes a positive outcome. Based on informal questioning of participants in previous experiments we find many claim to hold an internal representation of the “winning stimulus” just prior to feedback. If this is typical, then a negative peak attributed to the FRN may in fact partly reflect a perceptual mismatch phenomenon. Because feedback stimuli in the present experiment could be any number in a sixty point range, we expected this to undermine any habitual representation of a discrete winning stimulus, and so prevent the opportunity for a mismatch component to be introduced for –RPEs specifically.

Our aim in this experiment was to isolate neural activity associated with RPEs. Yeung et al. (2005) have observed that instrumental learning experiments, in which participants believe they are performing a feedback-guided learnable task, typically confound RPEs with error signals, that is, with non-economic judgements of whether an error was committed. To avoid this, participants in the present experiment were explicitly told that outcomes were unrelated to their key presses. Yeung et al. showed that while the FRN is reduced under such conditions it is still present, and a number of other studies have been able to demonstrate an FRN in conditions where participants were aware they could exert no control over an outcome (Donkers et al., 2005; Donkers & van Boxtel, 2005; Holroyd, Krigolson, & Lee, 2011; Marco-Pallares et al., 2010; Potts et al., 2006; Yeung & Sanfey, 2004). We nevertheless used a task that superficially resembled an instrumental learning task, in which participants selected from a choice of icons, to facilitate comparison with the literature.

In summary, our motivation in the present experiment was an attempt to separate components coding for + or – RPEs by using a data-driven methodology that removed the need for fixed-interval quantification, and the incumbent risk of combining

components within a single analysis window. The rationale for the design was that it should avoid known confounds and that it should be able to describe the relationship between RPE utility and voltage independently for +RPEs and –RPEs and map these relationships onto the canonical response function templates shown in Figure 13: describing the data as correlation coefficients served this purpose well. Based on the extensive FRN literature, we hypothesised that RPE utility encoding would be present (i.e. not just a UPE size response), but we had no prior hypothesis regarding which of the forms 1a – c this would take, nor when it would occur.

3.3 Methods

3.3.1 Participants

Sixty two undergraduates (9 left-handed, 22 male) participated for course credit and the opportunity to win a small sum of money. Data from 7 participants were rejected (five for equipment failure, two for eye blink artefacts on over 50% of trials).

3.3.2 Task Rationale

Prediction errors were manipulated using reward magnitude. Participants undertook gambles in separate gain and loss domain blocks, hoping in gain domain blocks to win money, and in loss domain blocks to avoid loss of money. A blocked design was used for the domain variable since the study sought only to control, not study, domain effects and it was believed that alternating domain on a trial by trial basis would confuse some participants and reduce FRNs generally as feedback stimulus - reward associations were continually being reversed.

On each trial, participants received feedback on the outcome in the form of a number in a sixty point range (23 to 82) denoting the points won (gain domain trials) or lost (loss domain trials). In each domain there were thus a range of thirty +RPE and thirty –RPE outcomes (centred around an EV of 52.5 , independent of whether points

were actually won or lost. The principal variable of interest was RPE utility, that is, a signed value corresponding to the difference between the actual points gained or lost on each trial and the average and therefore expected value of 52.5 points. Participants were not explicitly informed that 52.5 was the expected value, in order to avoid the possibility they might impose a categorical good vs. bad discrimination at this point. It was anticipated that exposure to the outcomes would fairly quickly allow the extraction of expected value by any neural component devoted to this process.

3.3.3 Task Procedure.

The experimental task was presented using E-Prime software and is summarized in Figure 14. Participants were shown a graphic depicting four symbols and selected one using a keypad. A fixation cross appeared (600–700 ms duration), followed by the points won or lost on that trial (700 ms duration) and then a blank screen (800 ms duration). Participants performed the task in 32 blocks of 60 trials each. In every block, each of the numbers 23 to 82 appeared once in a random order. For half the participants the first sixteen blocks were gain domain blocks, and the second sixteen were loss domain blocks, for the other half this pattern was reversed. Participants were verbally informed at the change point. At the end of each block, participants were shown a number that was ostensibly, though was not in fact, the sum of all the numbers that had been displayed in that block. If this fictional number was greater than 3,000, participants either won £2 (gain domain block), or lost £2 (loss domain block); if the number was below 3,000 there was no effect. In fact the value of this fictional number was predetermined to be over 3,000 half the time in each domain for a net gain of £0.00 across the experiment for this aspect of the task. Participants were told (truthfully) that their key-presses were unrelated to the outcome of the trial, and were told (untruthfully) that each trial was determined by a random number generator set to return an average points value of 3,000 across each block (the value was in fact fixed at 3150, the sum of

the values 23 to 82). It was explained to participants (again untruthfully) that each run of the random number generator was independent, and so it was unknown how much money they might make or lose on this aspect of the task: this was to encourage interest in the feedback.

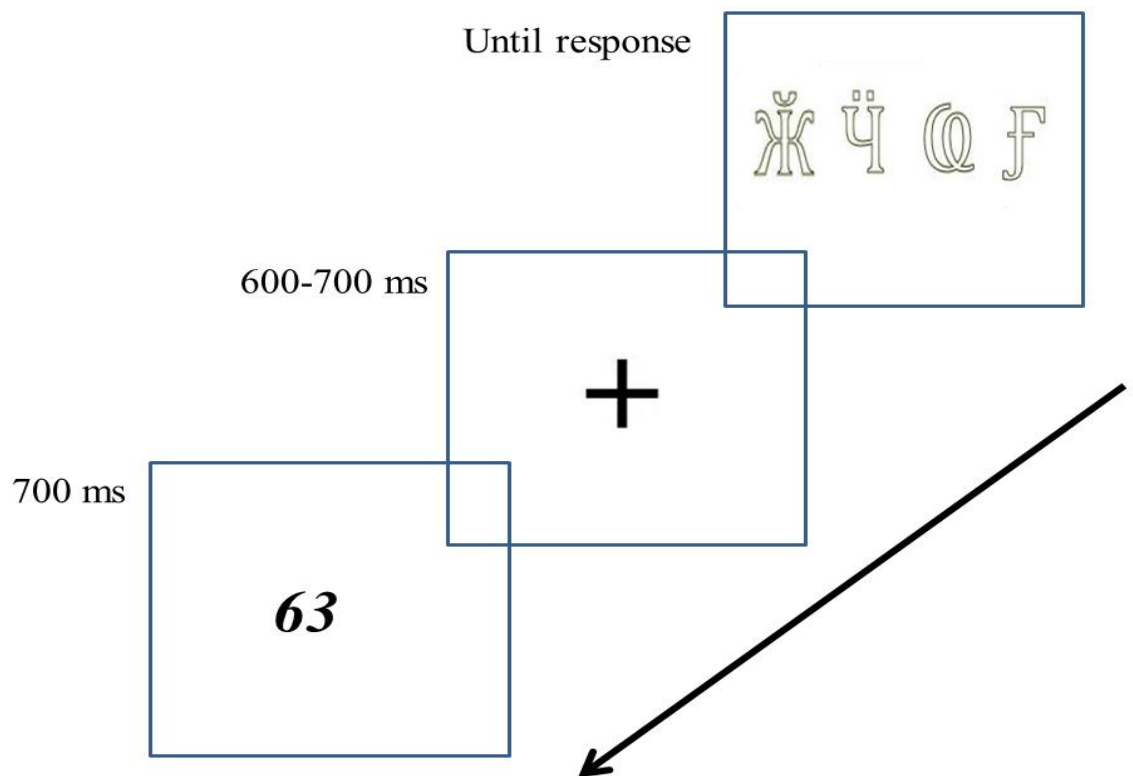


Figure 14 Summary of a single trial of the task procedure

In order to motivate participants to pay attention, ten percent of trials (192 trials) were followed by a probe in which participants had 1 s to make a key press to indicate what the number they had just seen was, versus a number that was one point higher or lower. Participants lost £0.01 with each failure. Since this could be done simply by stimulus matching against the just departed stimulus, i.e. without processing magnitude, a further probe was included. At the end of each block, before being told the points total

for that block, participants estimated whether the points total had exceeded or fallen short of 3,000 and were awarded £0.20 each time this matched the reported sum. Since the reported sum was, unbeknownst to the participants, fictitious and randomly determined, participants could only perform at chance on this question (50% average success rate): it was included merely to motivate attention to the points awarded on each trial of the block that preceded the question. Participants played 1,920 trials in total, with a 30s break between each of the 32 blocks. Total earnings for the experiment averaged £2.91 per person (approx. \$4.50).

At the conclusion of the experiment a check on participant preferences for outcomes was made. Participants were asked to rate how happy they were to see the outcomes 28, 38, 48, 58, 68 and 78 in each of the two domains by using a mouse to place a mark on a continuous rating scale labelled at one end as “very unhappy” and at the other end “very happy”. This process was repeated two further times and an average taken.

3.3.4 EEG Recording

EEG data were collected from 11 actively amplified Ag/AgCl electrodes (actiCAP, Brain Products, Gilching, Germany) mounted on an elastic cap. The electrodes were Fz, FCz, Cz, CPz, Pz, F3, F4, FC3, FC4, FP1, FP2. Electrodes were referenced to the left mastoid and re-referenced off-line to the average of left and right mastoid activity. Vertical eye movement was monitored by electrodes FP1 and FP2 and a right suborbital electrode, and horizontal eye movement was monitored using an electrode on the right external canthus. Electrode impedances were kept below 20 k Ω . EEGs were amplified using a BrainAmp amplifier (Brain Products), continuously sampled at 500 Hz, and filtered offline with a band-pass filter from 2 to 30 Hz designed to remove P3 effects. ERPs were computed by averaging artifact-free EEGs (~86% = 1650 trials). EEGs were rejected if eye movement electrodes showed a voltage change

exceeding 75 $\mu\text{V}/200\text{ ms}$ or if any midline site showed either a voltage change exceeding 20 $\mu\text{V}/\text{ms}$ or exceeded a value of $\pm 100\text{ }\mu\text{V}$ relative to baseline. EEGs were time-locked to 200 ms before the onset of the feedback to 700 ms afterward, and then were baseline-corrected using the interval -100–0 ms.

3.3.5 EEG Analysis.

While feedback consisted of a single number, this could correspond to quite different quantities depending on the economic terms by which it was evaluated. These are laid out in Table 1. In order to analyze RPE utility effects, stimulus values were recoded to the RPE utilities they represented as indicated in the table. This resulted in a range of sixty utility values running from -29.5 to +29.5 that was independent of whether the outcome constituted an actual loss or gain. The analysis was performed using correlations, and these were performed separately for –RPEs and +RPEs and individually for each participant. (A correlational analysis of this sort was inappropriate in the previous chapter because RPEs were manipulated as a simple +RPE /–RPE dichotomy) A Pearson correlation coefficient between voltage and RPE utility was calculated at each time point. Data points in this correlation corresponded to individual trials. Since each participant saw an average of 1650 trials, of which half were +RPEs and half were –RPEs, correlations were obtained from an average of 825 sample points. While the RPE utility was fixed for a given trial, because the voltage varied over time, the correlation coefficient on each trial therefore also varied by time. The correlation coefficients were plotted against time to produce a figure which was analogous to conventional ERP plots but which showed the *strength of RPE encoding*, derived from the full range of utilities experienced, rather than *actual voltage* for a given bracket of RPE utility (e.g. high) as would be used in a factorial design. The interpretation of such a figure is straightforward: where the waveform is at baseline there is no effect of the variable (RPE utility) on voltage, where there are deviations from baseline this indicates

a relationship, suggesting that RPEs are being coded by voltage. In this respect, such a correlational waveform can be interpreted as though it was in fact a traditional difference wave. Points on this waveform showing significant deviations could then be found by conducting a one sample t -test on the values of r (relative to an expected value of 0 under the null hypothesis) at each time point over the 55 participants. To facilitate subsequent Monte Carlo simulations, the one sample t -test was implemented as a paired samples t -test comparing a column of 55 observed values of r against a column of 55 expected values set to zero.

Table 1 Properties of feedback under different economic dimensions

Domain	<i>Feedback</i>			
	Stimulus	RPE valence	RPE utility	UPE size
Gain	23	−RPE	−29.5	29.5
	24	−RPE	−28.5	28.5
	...			
	52	−RPE	−.5	.5
	53	+RPE	.5	.5
	...			
	81	+RPE	28.5	28.5
	82	+RPE	29.5	29.5
Loss	23	+RPE	29.5	29.5
	24	+RPE	28.5	28.5
	...			
	52	+RPE	.5	.5
	53	−RPE	−.5	.5
	...			
	81	−RPE	−28.5	28.5
	82	−RPE	−29.5	29.5

The multiple comparisons resulting from the analysis of the whole waveform were addressed using a method based on the widely used cluster randomisation procedure of Maris and Oostenveld (2007). This procedure allows analysis of the entire

ERP waveform without incurring the excess conservatism of a strict Bonferroni correction for each time point analysed. It achieves this by recognising that because voltages are strongly correlated at adjacent time points the effective number of comparisons being made is much lower than the number of sample points in the waveform. While the technique's use in EEG is usually applied to voltage values it is equally appropriate for values of Pearson's r : it is simply concerned to test whether scores are significantly different from a test value of zero

In the first step of the procedure a one sample t -test on the values of r was performed at each time point and in each electrode channel in the manner described above, and used to identify significant ($p < .05$) t -values. Because of the gradual growth and decay of the correlation coefficients both over time and space, these significant t -values also occurred in clusters of time points and electrode sites. Clusters were identified by finding significant t -values that were contiguous in time or space (adjacent time points for the same electrode or electrodes within 4mm of each other and at the same time point). Only clusters containing eight or more samples (i.e. 16 ms) were considered for analysis. For each such cluster, a cluster-level t -value was calculated as the sum of all single sample t -values within the cluster. Analysis thereafter was based on these clusters and their associated cluster level t -value, rather than the individual (and highly non-independent) t -values derived from time x electrode points, reducing the multiple comparisons to a manageable number.

Since cluster level t -values could not be tested for significance against a standard t distribution, in step two of the procedure, the significance of each cluster was calculated by comparing its cluster-level t -value to a Monte Carlo generated distribution of cluster level t -values in the interval occupied by the cluster. This was generated from 50,000 datasets that corresponded to the null hypothesis. To create such a dataset, for each time point within the interval under consideration, the 55 observed r values were

paired with 55 expected values of 0, but this time observed and expected values were switched for a randomly determined number of participants prior to running the *t*-test, ensuring a *t*-statistic based on data corresponding to the null hypothesis. The process was repeated for all time points in the interval under simulation and the *t* values were summed to create a cluster level *t* statistic under the null hypothesis. Having generated 50,000 such cluster level *t* statistics, the *p* value of the actual observed cluster found in step 1 was calculated as the proportion of the randomisation null distribution that exceeded its cluster-level *t* statistic. A Bonferroni correction was then made such that alpha was set to .025 divided by the number of clusters found to be significant at step one. The process was then repeated for the next cluster identified at step one..

3.4 Results

3.4.1 Behavioural Data

Participants were found to answer the probe question correctly on 73.87% of probed trials, with no significant difference in accuracy between loss domain blocks and gain domain blocks (paired samples *t* test: $t_{54} < 1$). Preference data collected at the end of the experiment showed a very high correlation between RPE utility and rating ($r = .91$, $p < .001$, $N = 660$). The relationship between RPE utility and rating was the same for both gain and loss domains as shown by very similar average beta values for participant-wise regression of rating against RPE utility (mean β for gain = 0.93, mean β for loss = 0.94) with a paired samples *t*-test on the beta values showing no significant difference ($t_{54} < 1$). These results indicate that participants attended to feedback and affectively responded to it an appropriate way. They also suggest that the domain effects on conscious preferences that were seen in Chapter 2 (greater sensitivity to outcome in the loss domain) could be removed by blocking gain and loss sections of the experiment separately.

Participants had been told that there was no link between their behaviour and the outcome of the trial. We checked to see if their behaviour was consistent with this belief. Although the outcome of trials was indeed random, this randomness led to some icons being more profitable on a given block than others and it is possible that participants modified their choices on this basis. On a participant/block basis, and after removing cases where icons were never or always chosen by a participant, there was indeed a correlation between the profitability of icons and the frequency with which they were chosen ($r = .04$, $p = .001$, $N = 6880$). We also examined whether the RPE utility on a trial affected the likelihood of switching to a different icon on the following trial and whether this depended on the RPE valence. While intuitively the dependent variable for this analysis would appear to be the decision to switch, and the independent variables to be the RPE utility and valence on the present trial (since these cause the switch), this would result in a binary dependent variable, to be avoided when using ANOVA. Thus, as ANOVA is indifferent to the temporal relationship of the variables we allocated switching behaviour on the next trial (switch vs. no switch) and RPE valence (+ vs. -) as independent variables, and RPE utility on the current trial as the dependent variable. The results showed that RPE utility was indeed significantly lower on trials which were followed by a switch of icons ($F_{1,105594} = 7.74$, $p < .001$, $\eta^2 < .001$). This utility difference was greater for -RPE trials (.29 points) than for +RPE trials (.03 points), producing a significant interaction term ($F_{1,105594} = 4.97$, $p = .026$, $\eta^2 < .001$). Thus despite our instructions, to some degree the participants appeared to regard the experiment as a reinforcement learning task, and furthermore showed different learning effects for rewards and punishment. These effects, while significant, were of negligible size however, as indicated by the small η^2 values. Moreover, the speed with which icons were chosen following the presentation of the icon choice array (interquartile range: 261–592 ms) suggests that for the most part icons were regarded as unimportant, that

button presses were simply used to elicit outcomes, and that only feedback was processed.

3.4.2 Electrophysiological Data

The scalp topography of valence effects, captured by a difference wave of – RPEs and +RPEs is shown in Figure 15. In keeping with the literature, this shows an early frontocentral negativity, shifting parietally at greater latencies. Simple waveforms (with no high pass filter) for +RPEs and –RPEs are provided, along with the associated difference wave at Fz. As anticipated however, the participants’ passive stance, and the absence of categorical winning and losing stimuli resulted in weak effects which would not necessarily be evident in a simple factorial comparison of +RPEs and –RPEs, so we now turn to the more powerful correlational analyses.

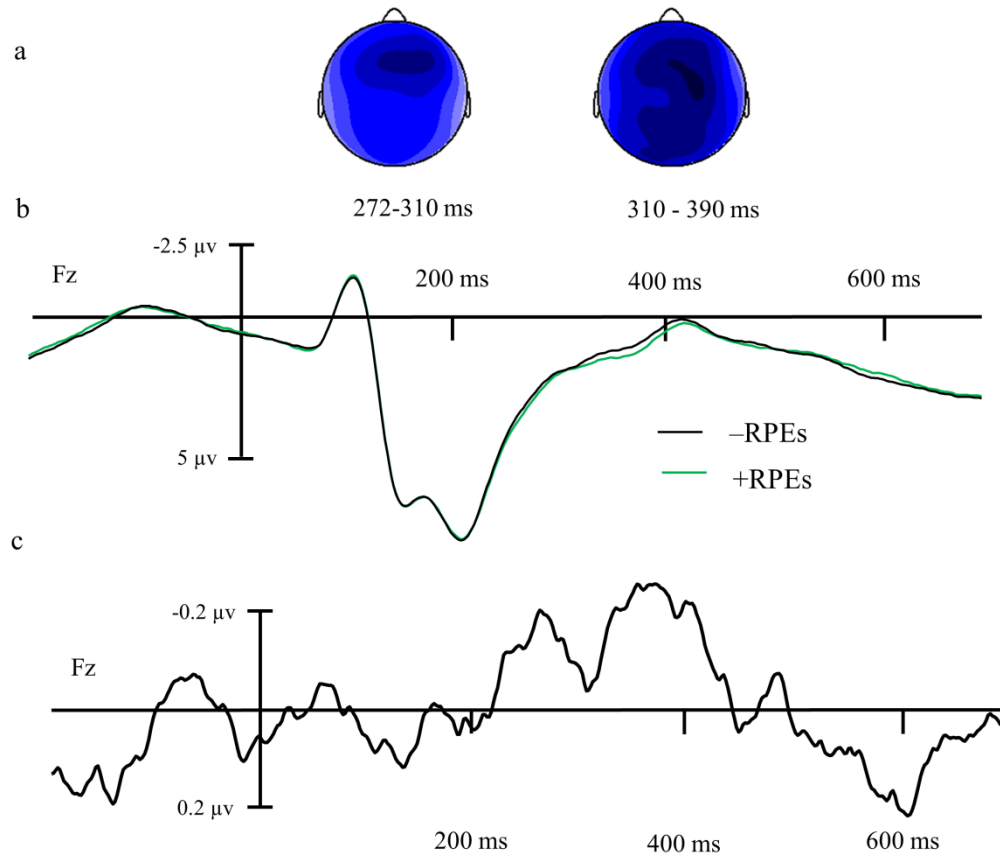


Figure 15 Effects of valence

Effects of valence, that is, the average of all +RPEs contrasted with all -RPEs. a) scalp topography of difference, b) grand averages at Fz, c) difference wave at Fz.

Figure 16a shows the grand average correlation coefficients of voltage and utility at three representative midline sites. In this figure, deviations from baseline indicate points in time at which voltage amplitude appears to code RPE utility. The approximate threshold for significance ($p < .05$) is shown; however this suffers from the multiple comparisons problem leading to an inflated possibility of Type I error. To correct for this, the Maris and Oostenveld cluster randomisation procedure described above was used, applied in the interval 100–700 ms and at Fz, FCz, Cz, CPz and Pz, and the surviving clusters of RPE-related activity in the waveform are shown in Figure 16b. These clusters are considered individually below. To aid visualisation of how Figure 16 would correspond to a standard time x voltage plot, readers may refer to Figure 17.

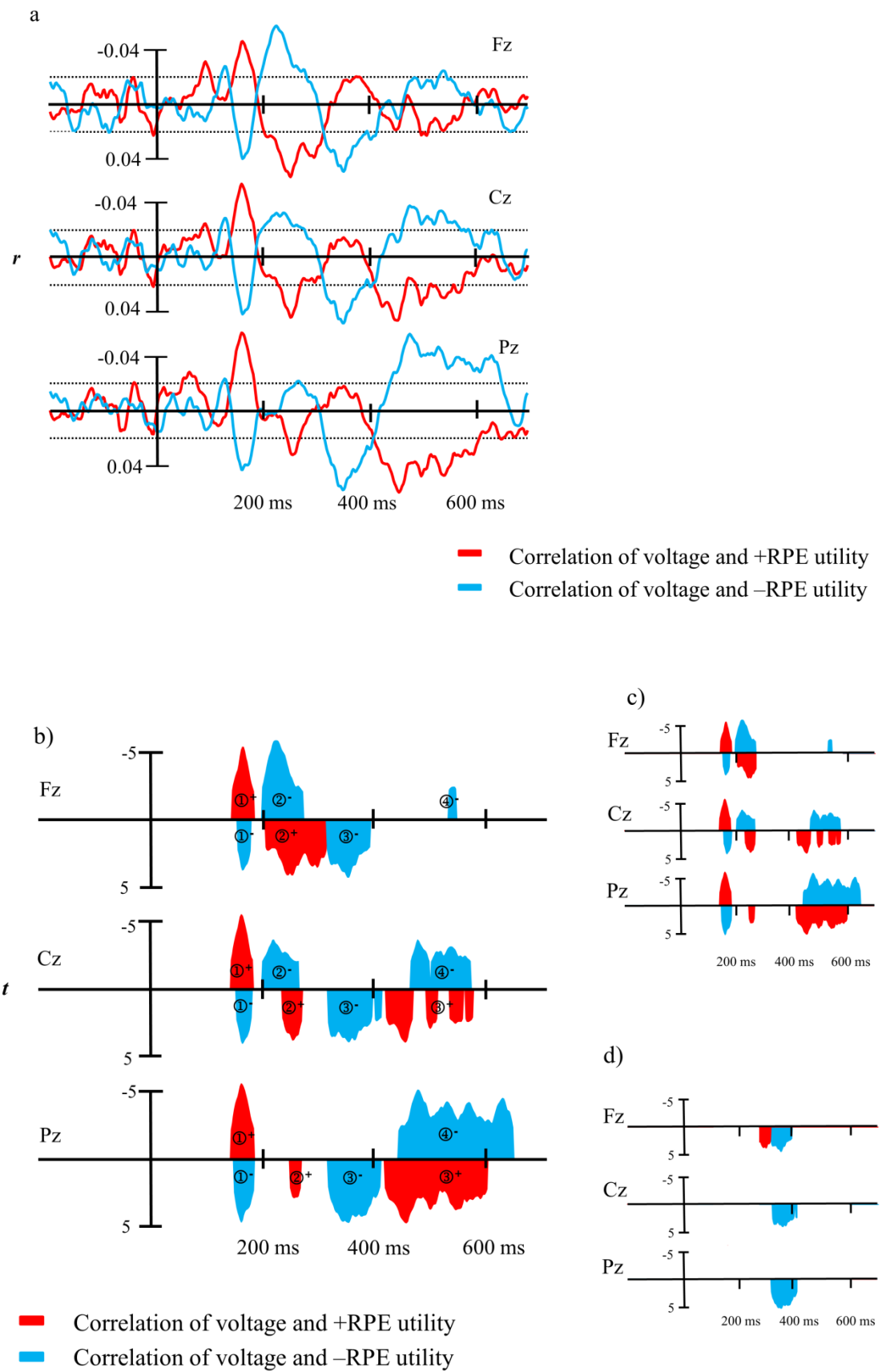
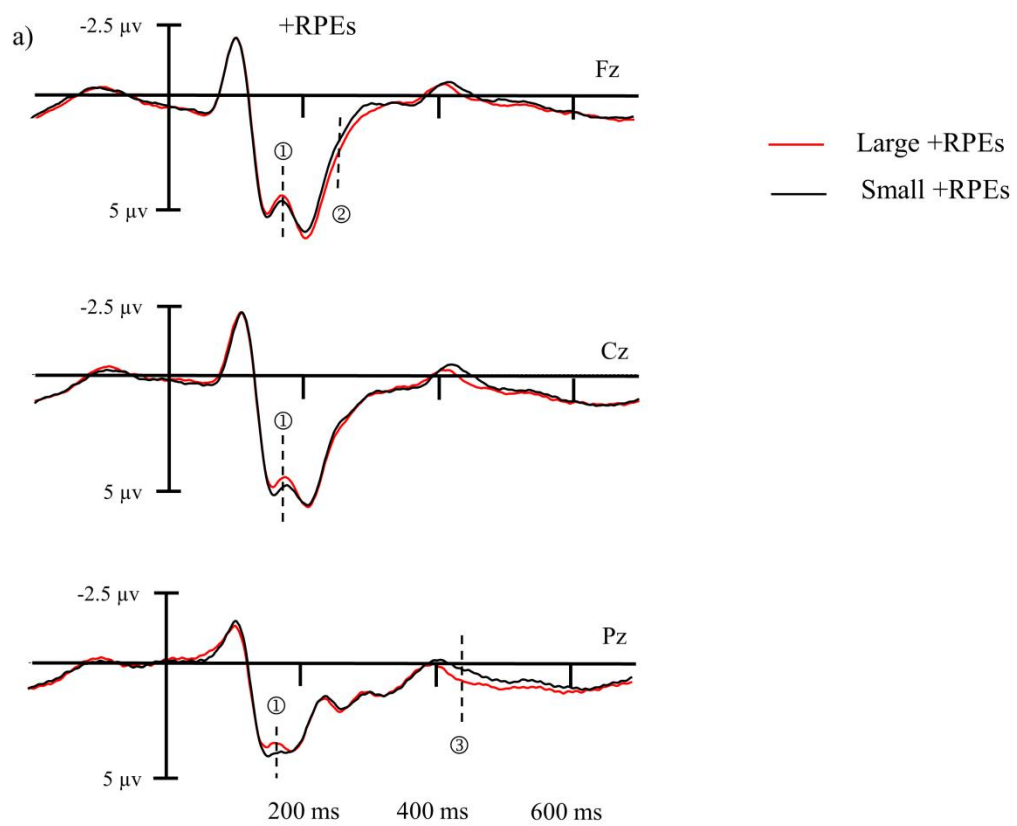


Figure 16 Correlational waveforms

Grand average of Pearson correlation coefficients of voltage and RPE utility. Dotted lines show approximate threshold for significance of a one sample t-test of correlation coefficients ($N = 55$, expected value of r under the null hypothesis is 0). b) Intervals of significant correlation between voltage and RPE utility after correction for multiple comparisons. t -values were obtained from a one sample t-test of correlation coefficients ($N = 55$, expected value of r under the null hypothesis is 0) at each time point and formed into clusters of significance for which a cluster-level t value was calculated as the sum of individual t -values. This was compared to a Monte Carlo simulated distribution generated under the null hypothesis ($r = 0$) to establish significance. Non-significant ($p > .05$) values of t have been set to zero. See text for further details c). Intervals coding for unsigned prediction error size, or salience. d) Intervals coding for RPE utility



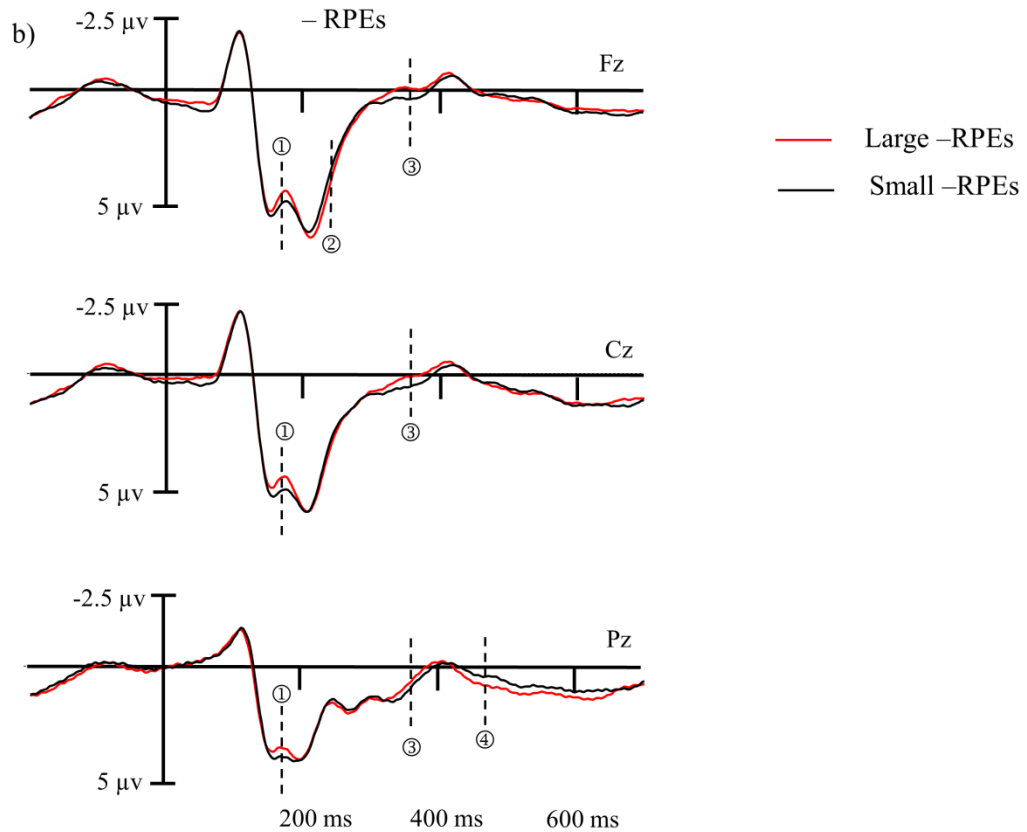


Figure 17 A factorial rendition of the experimental design

A factorial rendition of the experimental design with the +RPE utilities bracketed into two categories, large and small, and -RPE utilities bracketed likewise. Salient outcomes (large +RPEs and large -RPEs) are shown in red. Points of the waveform in Figure 16 that show strong correlations of RPE utility and voltage should, in these figures, thus be points at which there is a difference between voltage for large and small RPEs. These are indicated with circled values that correspond to the clusters in Figure 16b

Analysis of the +RPEs revealed three clusters of activity that were significant at the .008 threshold set by Bonferroni correction: cluster 1⁺ occurring at all midline sites from ~140–180 ms ($p = .000007$), cluster 2⁺, for which the largest temporal response was at Fz from 204–312 ms ($p = .0003$) and cluster 3⁺ for which the largest temporal response was at Pz from 418–600 ms ($p = .000002$).

Analysis of the –RPEs revealed four clusters of activity that were significant at the .004 threshold set by Bonferroni correction: cluster 1[–] occurring at all midline sites from ~145–180 ms ($p = .0001$), cluster 2[–], for which the largest temporal response was at Fz from 198–272 ms ($p = .00002$), cluster 3[–], occurring at all midline sites from ~310–390 ms ($p < .000002$) and cluster 4[–], for which the largest temporal response was at Pz from 442–648 ms ($p = .000006$).

3.4.3 Discrimination of Salience and Utility effects.

An important objective of the experiment was to distinguish intervals in the post-feedback waveform which appeared to code for the utility of +RPEs and –RPEs from those merely coding for the size of UPEs. Since UPE size and RPE utility are perfectly positively correlated for +RPEs (the bigger the prediction error the higher the utility) and perfectly negatively correlated for –RPEs (the bigger the prediction error, the lower the utility), intervals in which a UPE size response occurs should show a response to RPE utility that is oppositely signed for +RPEs and –RPEs. Figure 16b allows identification of these. From ~145–180 ms, across the midline broadly, a UPE size response appears to occur, as indicated by correlations between voltage and utility which are oppositely signed for +RPEs and –RPEs. The negative sign of the correlation for +RPEs indicates that in the original waveforms increased utility for good outcomes was associated with increased negativity of voltage. In contrast, the positive sign of the

correlation for –RPEs indicates that in the case of bad outcomes *decreased utility* (i.e. large –RPEs) was associated with negativity of voltage. Thus large prediction errors, regardless of their sign, induced a voltage- negativity in this interval. A second UPE size response was seen from ~200–270 ms, strongest at Fz. The reversal of the correlation signs relative to the preceding UPE size response indicates that here large prediction errors, regardless of their sign, were associated with positivity of voltage. Note that responsiveness to +RPEs alone persisted beyond this interval up to 312 ms, an effect we consider below. From ~310–390 ms, across all midline sites, there was a significant correlation between voltage and the utility of –RPEs only. The positive sign of the correlation indicates that positivity of voltage was associated with increased utility (i.e. small –RPEs showed more positive voltages than large –RPEs). Finally, from ~400–640 ms, and most pronounced at Pz, another UPE size response was seen, with positivity of voltage associated with large prediction errors. Figure 16c shows the incidence of UPE size responses alone, with utility effects removed.

A considerable advantage of the cluster randomisation technique is that it does not require any assumptions about the timing of components, instead locating *all* intervals of significant activity while nevertheless avoiding increased Type 1 error. Notwithstanding this absence of *a priori* stipulations about where effects should be measured, it is still appropriate to interpret its results in the light of theoretical expectations. One such expectation is that a response to RPE utility will occur in the interval 200–350 ms. In keeping with this expectation, Figure 16b suggests an encoding of +RPEs from 272–310 ms at Fz. However, the cluster randomisation technique has combined this effect with the earlier UPE size coding. This means that a statistical demonstration of coding of +RPE utility is wanting, since while the transition from UPE size to +RPE utility coding at 272 ms is visually compelling, the significance of the +RPE coding depends to an unknown degree on conglomeration with the earlier UPE

size effect. To establish the reality of the +RPE encoding in the 272–310 ms interval, the cluster randomisation procedure was run only in the interval 272–700 ms. The cluster remained significant at the .0125 threshold set by Bonferroni correction: $p = .002$). Figure 16d shows the incidence of utility responses alone, with UPE size effects removed.

The use of Pearson correlation coefficients to examine RPE encoding presumes that the relationship between voltage and utility is linear, and this assumption was made to simplify the analysis and its exposition. Having identified intervals of encoding of +RPEs from 272–310 ms and –RPEs from 310–390 ms under this linear assumption, we examined the nature of the relationship by plotting grand average voltages at each of the sixty levels of utility. Figure 18a reveals a linear relationship between voltage and utility for both +RPEs and –RPEs. Figure 18b provides the corresponding plot for the salience effects in the interval 200–270 ms. For comparative purposes, the x axis in this second plot has been left as RPE utility, however the V-shaped function clearly suggests that it is UPE size that is being coded, and with respect to this variable, the function is once again linear.

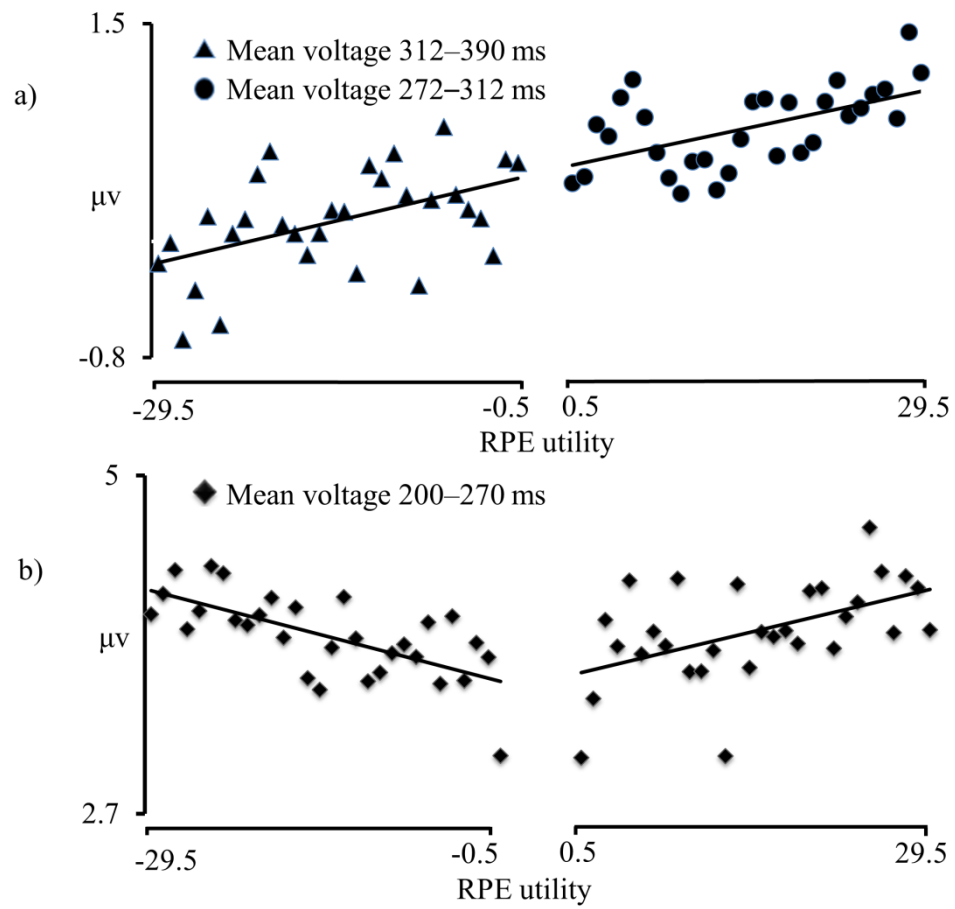


Figure 18 Grand average of voltages at each level of RPE utility

Grand average of voltages at each level of RPE utility at Fz for a) +RPEs in the interval 272–312 ms and –RPEs in the interval 312–390 ms, and b) all prediction errors in the interval 200–270 ms.

3.4.4 Domain Effects.

The study was concerned to control rather than study domain effects. This control was achieved by presenting separate gain and loss domain blocks, in which +RPEs and –RPEs were represented equally often. As well as controlling domain effects, the blocking was expected to reduce domain effects compared to Chapter 2, in which domain was deliberately foregrounded by switching domain from trial to trial at the participants’ own direction. Nevertheless we considered it worthwhile seeing whether domain effects persisted in a blocked design, since this is of methodological concern given that most FRN experiments present gambles in the gain domain only, or confound domain and valence by presenting mixed gambles. Figure 19 shows the same correlation waveforms for +RPEs and –RPEs that were shown in Figure 16a, but broken down into the two domains. It can be seen in Figure 19a that in the interval for +RPE encoding (272–310 ms), gain domain gambles generate a stronger signal, while in Figure 19b, in the interval of –RPE encoding (310 – 390 ms) loss domain gambles create the stronger signal. As such, it appears that unusually large losses and large gains create a stronger signal than unusually small losses or small gains. This was investigated with a 2 x 2 (domain x valence) ANOVA on the average value of r , using the interval of 272–310 ms for +RPEs and 310–390 ms for –RPEs. This showed a non-significant interaction ($F_{1,54} = 2.92$ $p = .09$) suggesting that domain does not affect the relative strength of +RPE and –RPE encoding. Overall, there was no main effect of domain ($F_{1,54} = 1.69$, $p = .20$). Thus it seems that blocking the domain variable removed its effects, and, furthermore, there is no evidence that an experiment which use only one domain (typically gain) would artificially increase the strength of +RPE or –RPE coding.

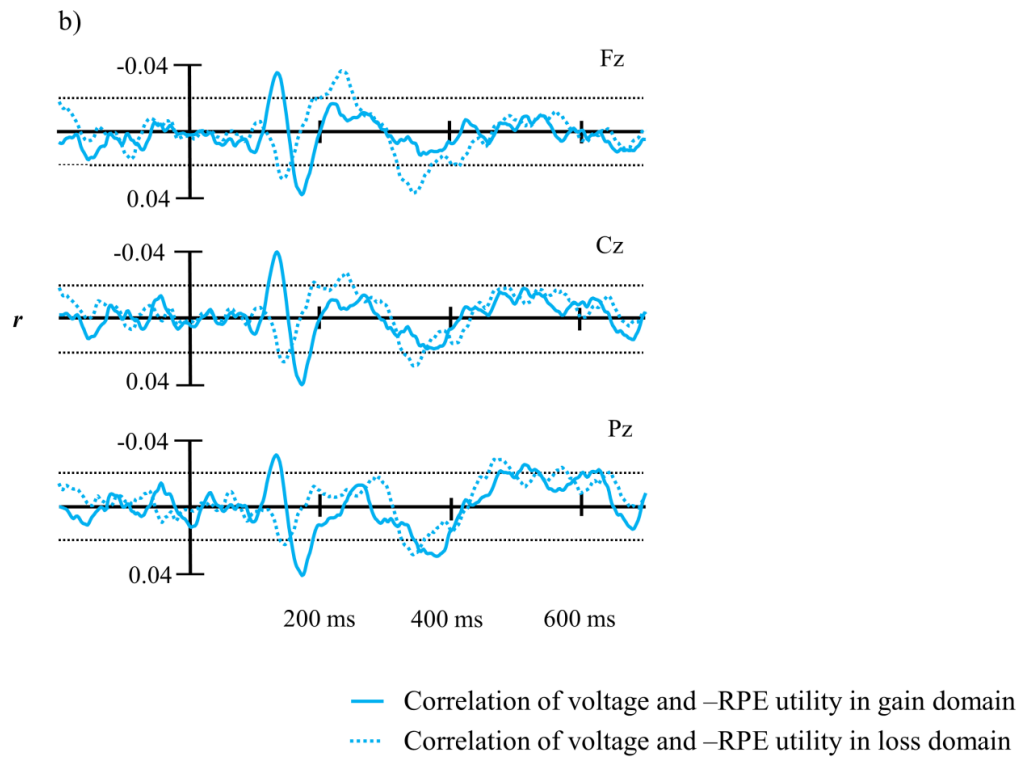
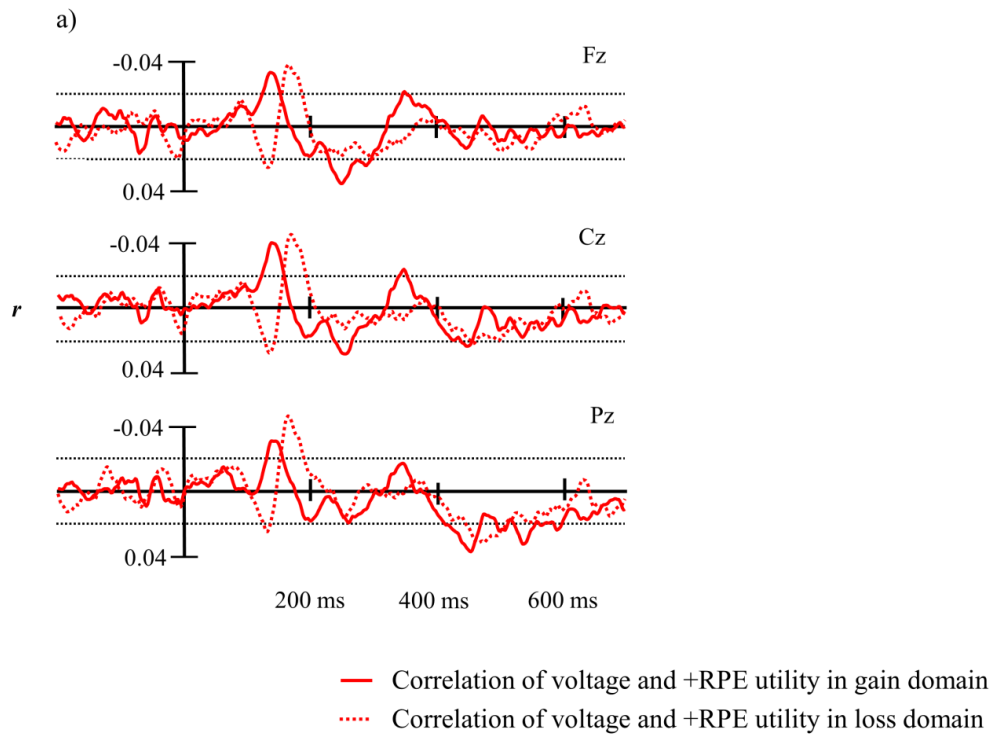


Figure 19 Correlational waveforms broken down by outcome domain

3.5 Discussion

An effective valuation system should be sensitive to both punishment and reward and so should make use of the quantitative information held in the utility of both +RPE and –RPEs. The present experiment investigated whether such sensitivity to both +RPEs and –RPEs was shown in mediofrontal ERPs. To do this, we initially identified intervals of the waveform responsive simply to UPE size, since this might have complicated previous interpretations of RPE utility effects. While early and late intervals suggested coding for UPE size, the study suggested that in the interval ~270–390 ms RPE utility, not UPE size was coded. In this interval, the mediofrontal waveform was responsive to the utility of both +RPEs and –RPEs. Importantly, increased utility in both kinds of RPE was associated with increased positivity of voltage. This is important because integrated coding of +RPEs and –RPEs as depicted in Fig 1a requires the relationship between utility and voltage to be same-signed for both RPE valences in order to establish a common currency of utility for reward and punishment.

It is worth noting here that the value of the correlation sign is not meaningful in itself and the implications would be unaltered if, in the interval 270–390 ms it was negative for both +RPEs and –RPEs. To draw an analogy with more familiar examples, while the sign of traditional components such as the N2 and P3 is clearly pre-eminent for those components' identification, it does not itself convey any information regarding the component's function. Furthermore, the fact that for both +RPEs and –RPEs the sign of the relationship between utility and voltage alternates across the waveform, as shown in Fig 4, is not remarkable and can be assumed to indicate sequential and

independent components responding to the same properties of feedback but with opposite polarity. A well-known existing example of this is the N2–P3 complex in which unexpected events produce an ERP showing an accentuated negative peak followed by an accentuated positive peak (Folstein and van Petten 2008). Indeed, an analogous effect can be seen in Figure 17, where large UPEs are associated with an accentuated negative peak at ~175ms and then an accentuated positive peak at ~210 ms and this is reflected in Figure 16 by a correlation sign switching at ~200 ms.

Returning to the observed utility effects, the correlation of utility and voltage for +RPEs at 272–310 ms is notable not for the actual sign of the correlation but simply for the absence of an oppositely signed correlation for –RPEs in the same interval; it is this absence that allows us to discount the possibility that this is a UPE size response, leaving an RPE utility response as the most probable alternative. The same inference applies to the correlation of utility and voltage for –RPEs at 310–390 ms, and collectively this forms the basis for our primary conclusion that both +RPE utility and –RPE utility are coded in the mediofrontal feedback locked waveform.

As noted above, the relevance of the correlation sign lies in the fact that increased utility was associated with increased positivity of voltage for both kinds of RPE. This suggests the possibility of a single RPE processing system capable of assigning a utility value to all outcomes, both good and bad, in a manner directly comparable to that used by formal reinforcement models. Moreover, the consecutive nature of the RPE utility signal, with –RPE coding following +RPE coding would appear to be consistent with one account of midbrain RPE generation. It has been proposed that the midbrain dopamine neurons that code +RPEs with phasic increases are limited in their ability to code –RPEs with phasic decreases because of their already low tonic firing rate, and may therefore code the utility of –RPEs with the *duration* of firing decrease (Bayer et al., 2007; Mileykovskiy & Morales, 2011). This would result

in a delay in the transmission of –RPEs to the neural generator of the FRN consistent with the effect shown here. Note however that the scalp topography of the response to +RPE and –RPE utility differed, with +RPEs most pronounced frontally, but –RPEs showing a broader distribution. While it is possible that there is a single frontal source for all RPE-related activity and then an additional source for –RPEs specifically, it is also possible that there is a distinct generator for each kind of RPE. While functionally there is no need for –RPEs and +RPEs to be processed in the same neural structure in order to effect reinforcement learning this does of course undermine the case for a single unitary RPE encoder.

The current study suggested the strength of coding of –RPE utility to be stronger than coding of +RPE utility as indicated by much more significant Monte Carlo p values. However, a recent review of the FRN by Walsh and Anderson (2012) found coding of +RPE utility to be more commonly observed than –RPE utility coding. One source of this asymmetry may well be the interval of measurement of the FRN, which shows wide variability across the literature, but often does not extend beyond 350 ms or even 300 ms, and thus could miss later –RPE effects. A second source concerns interference from the P3. While this study used a 2 Hz filter in an attempt to mitigate P3 effects, many FRN studies do not. Since the P3 is typically more positive for unexpected and high magnitude outcomes regardless of valence (San Martin, 2012), that is, it is sensitive to UPE size, a late +RPE signal would be augmented by this UPE effect while a late –RPE signal would be diminished. This constitutes a specific example of the distorting effects that components encoding UPE size can have on interpreting utility effects in components that they overlie.

Two other experiments have investigated the ERP response to parametrically manipulated +RPEs and –RPEs. Talmi, Atkinson, and El-Deredy (2013) explicitly looked for an integrated RPE utility coding across +RPEs and –RPEs and reported

failure in this regard, instead finding a UPE size response. This is likely a consequence of the interval chosen however, 200–300 ms, where this study also found a strong UPE size signal. Using concurrent MEG across a wider time range however, the authors found an integrated utility signal at 320 ms, close to the 310 ms point in the present experiment where +RPE utility coding switches to –RPE utility coding

Pedroni, Langer, Koenig, Allemand, and Jancke (2011) also investigated +RPEs and –RPEs and found consecutive responses at similar latencies to this study: +RPE utility coding at 290–310 ms and –RPE utility coding at 360–380 ms. An important difference however was that that study found the correlation between voltage and –RPE utility to be negative, not positive. A key difference between the study and that of ours and Talmi et al. (2012) is that in Pedroni et al.’s, –RPEs were achieved by omission of reward. Esber and Haselgrove’s (2011) claim of an inherent lowered salience of non-rewards has already been noted. Furthermore, single cell work has suggested the possibility of separate systems for coding reward delivery and reward omission, regardless of the utility that either eventuality comprises. In particular, Joshua, Adler, Mitelman, Vaadia, and Bergman (2008) have found midbrain dopamine neurons that increased their response to aversive outcomes but decreased it to omitted rewards, despite both these outcomes constituting –RPEs. Thus the manner in which –RPEs were generated may explain the reversal of the sign observed in Pedroni et al.’s study with respect to that of ours and Talmi et al.’s.

The utility effects at 272–390 ms occurred between what appeared to be strong responses to UPE size occurring in the intervals ~145–180 ms, ~200–270 ms and ~400–640 ms. These intervals were indicated by significant oppositely signed correlations of RPE utility with voltage, represented by the inverted waveforms of +RPEs and –RPEs in Figure 16. The UPE size effect at 200–270 ms is consistent with other demonstrations of strong UPE size effects in this interval (Hauser, et al., 2014; Talmi, et al., 2013) and

fMRI work showing such a signal (Metereau & Dreher, 2013; Rutledge, Dean, Caplin, & Glimcher, 2010). The sustained strong UPE size effect at Pz running from 400 to 640 ms corresponds spatiotemporally to the slow wave component, shown by Foti et al. (2011b) to be responsive to UPE size rather than RPE utility.

The demonstrated UPE size effects, especially in the 200–270 ms interval, appear to vindicate the concern that FRN measurements are susceptible to UPE size contamination. The direction of the UPE size effect, greater positivity for larger prediction errors, can be expected to increase voltage differences for +RPEs of different size and decrease them for –RPEs of different size, leading to an apparent preferential sensitivity of the FRN to +RPEs. As it happens, this study suggests that, in the interval in which the FRN is typically measured, the RPE response is indeed largely driven by +RPEs, however it is possible that UPE effects have been mistaken for RPE effects in the earlier portions of this interval.

The UPE size effect is also of interest in its own right, as the same effects have been seen in single cell recordings (Matsumoto, Matsumoto, Abe, & Tanaka, 2007) and in fMRI meta-analyses (Bartra, McGuire, & Kable, 2013), described as indicating motivational salience (Bromberg-Martin, et al., 2010). As large UPEs were as common as small ones in this study, the “salience” indexed by this UPE size response does not reflect novelty of the stimulus itself, but rather the notability of the appearance of a value that is an outlier with respect to a current estimate of a distribution. Such comparisons, and the updated expectations they produce, are central to Bayesian models of optimal foraging. They might, furthermore, benefit from independence of RPE circuitry: an animal that had eaten to satiation, for example, would still do well to note an unusually large source of food.

As we note earlier, an apparently unambiguous response function can still be an artefact synthesised from two components with quite different response functions where

overlap is complete. It is possible that an apparent UPE size response is in fact the aggregate effect of separate populations of neurons coding +RPE and –RPE utility. In this scheme, one population codes +RPE utility as in Figure 13a, and another codes –RPE utility as in Figure 13b, but with a reflected response function, that is with *increases* of activation for large –RPEs rather than decreases. Such response functions have been found in single cells (Asaad & Eskandar, 2011; Matsumoto et al., 2007) often in close proximity in the primate ventromedial prefrontal cortex and anterior cingulate cortex (Kennerley, Dahmubed, Lara, & Wallis, 2009; Monosov & Hikosaka, 2012; Quilodran, Rothe, & Procyk, 2008). This makes their separation by EEG and fMRI impractical, leaving the question unresolved at present of whether an apparent UPE size signal might in reality reflect an aggregate of two utility signals occurring earlier than the FRN, and which might in fact be involved in its generation. This interpretation is less parsimonious on face value, partly because it invokes two underlying mechanisms rather than one, but also because neural discrimination between large –RPEs and large +RPEs then requires a following neural integrator receiving excitation from one source and inhibition from the other in order to generate the required integrated utility function in Figure 13c. None of these possibilities are neurally implausible however, and the application of techniques other than EEG will be required to resolve these two interpretations of the UPE size signal that are both equally consistent with the data.

Finally, it is worth discussing the consequences of using a correlational waveform as the basis for analysis rather than a standard voltage-based waveform. It should be noted first that the correlation coefficients are small simply because they are based upon single trial ERP voltages, which are inherently noisy. In interpreting these correlations, the *p* value is more relevant than the *r* value because the latter is greatly affected by the amount of averaging prior to running the correlation. For example, take a representative correlation, $r = .04$, found on the +RPE correlational waveform of

Participant #6 at 306 ms. This value of r is based on ~825 individual data points of voltage plotted against thirty levels of RPE, which reveals a great deal of scatter between the points. By averaging voltage at each RPE level we are left with 30 data points, a much reduced scatter, and now a calculated $r = 0.21$. Reducing this to three bins of low, mid and high RPE gives a reported $r = .52$. This final level of averaging better approximates to the sort of comparison and effect size that is seen in standard factorial designs.

The question of peaks should also be addressed. FRN studies often, but by no means always, produce a waveform with an N2 peak superimposed on an ongoing positivity running from ~100–400 ms. Opinion varies as to the value of using this peak to identify the FRN: while some studies use its amplitude, or its amplitude relative to preceding or following adjacent peaks as a measure of the FRN, others ignore the N2 peak, instead using a simple mean amplitude measure across a set interval, or a measure based on a difference wave of good and bad outcome waveforms (either peak or mean amplitude). While we implemented such a peak to peak measure in Chapter 2 as part of a battery of tests, it is our opinion that the N2 peak of a single (i.e. undifferenced) waveform is not a reliable guide to either the amplitude or latency of the FRN component. The theoretical basis for this position has been clearly stated by Luck (2005), which is that peaks are not equivalent to components because each peak in a single waveform represents the summed effect of many components. A further important methodological objection to use of the N2 peak is that it may be absent, particularly for large +RPEs, and setting peak values to zero in such cases (e.g. Holroyd et al., 2003) produces a floor effect that distorts the comparison of +RPEs and –RPEs. As argued by Luck (2005), components are generally better described by *experimental effect*. This has traditionally been achieved by differencing pairs of waves drawn from different levels of an independent variable. While correlational waveforms are simpler

and more powerful in cases where the independent variable is continuous, they achieve exactly the same end and should produce qualitatively similar results. In comparison, single waveforms, taken in isolation, can produce quite different and possibly misleading phenomena. As a case in point, single waveforms for the present experiment (see Figure 17) show a negative deflection at ~175 ms. However, this is incidental to determining the latency of the FRN, since there is no effect of valence at this point and sensitivity to valence is by definition a property of the FRN component. In contrast, a UPE size effect does occur close to this negative peak, though of course that need not be so, a peak might occur with equal amplitude for large and small UPEs.

To conclude, this experiment investigated whether a mediofrontal ERP response existed to both +RPE and –RPE utility. The experiment controlled confounds that have not previously been controlled. Rather than using a standard factorial design, RPE size was manipulated parametrically, using a powerful correlational analysis. A cluster randomisation based correction for multiple comparisons allowed us to investigate the responsiveness of the waveform beyond the window typically used to assess the FRN without risk of false positives. Our results suggest that mediofrontal ERPs are responsive to the utility of both –RPEs and +RPEs. Insofar as the task was based on passive observation of outcomes, it resembled those tasks used in single cell studies of RPEs and, as such, its results strengthen the proposed link between the midbrain dopamine system and the FRN. The asynchrony of the response to +RPEs and –RPEs is also consistent with the behaviour of midbrain dopamine neurons as described by Bayer and Glimcher (2005). However, while they characterise the single cell literature, such passive tasks are not representative of the FRN literature, which relies more on instrumental learning tasks, to which we turn in Chapter 4.

4 An Examination of Sensitivity to Positive, Negative and Unsigned Prediction Errors with Principal Components

Analysis

The work in this chapter is strongly based on a paper currently under review (Sambrook and Goslin, under revision)

4.1 Chapter Abstract

Models of reinforcement learning represent reward and punishment in terms of RPEs, quantitative signed terms describing the degree to which outcomes are better than expected (+RPEs) or worse (–RPEs). An effective reinforcement learning system should be capable of representing the size of RPEs of both valences to allow the computation of net worth. The FRN has been claimed to neurally encode an RPE. An outstanding question however, is whether the FRN is sensitive to the size of both +RPEs and –RPEs. Previous attempts to answer this question have examined the simple effects of RPE size for +RPEs and –RPEs separately. However, this methodology can be compromised by overlap from components coding unsigned prediction errors, or “salience”, which are sensitive to the absolute size of a prediction error but not its valence. In our study, +RPEs and –RPEs were parametrically modulated using both reward likelihood and magnitude, with principal components analysis used to separate out overlying components. This revealed a single RPE encoding component responsive to the size of +RPEs at ~330 ms, and a number of other components responsive to unsigned prediction error size. No component sensitive to –RPE size was found.

4.2 Introduction

A key concept of reinforcement learning is that it is driven by both reward and punishment. When rewarded, actions are more likely to be repeated, when punished, less likely. While reward and punishment might appear qualitatively very different (e.g. food vs. electric shock) reinforcement learning models reconcile them with the underlying principle of the RPE, a numerical signed term describing the value of an outcome relative to its expected value. +RPEs indicate better than expected outcomes (i.e. rewards) while -RPEs indicate worse than expected outcomes (i.e. punishments). Turning rewards and punishments into numbers, differing only by their sign, makes them commensurable and allows calculation of the net value of a course of action that will incur both rewards and punishments, producing powerful reinforcement learning algorithms. However, although potentially powerful, is this “integrated coding” the approach used by the human brain, or are reward and punishment dissociated into separate systems?

The current evidence leans towards dissociation. For example, fMRI meta-analyses suggest that brain areas that code for both rewards and punishments are the exception rather than the norm (Bartra et al., 2013; Garrison, Erdeniz, & Done, 2013). Similarly, single cell studies show that neurons that fire in response to reward or punishment only are markedly more common than those that raise their firing in response to rewards and reduce it in response to punishments (Kobayashi et al., 2006). Furthermore, it has been argued that the dissociation of reward and punishment may be neuro-chemically instantiated, with dopamine coding reward and serotonin coding punishment (Daw et al., 2002).

The FRN has been claimed to represent an RPE used in reinforcement learning insofar as it shows a relative negativity for -RPEs compared to +RPEs. Such discrimination of RPE valence is clearly necessary for an RPE encoder, and indeed, an

axiomatic model of RPEs (Caplin & Dean, 2008) states this as its first axiom of RPE encoding. The second axiom in this model requires that changes in the FRN's voltage in response to increasing RPE size should not be in the same polarity for both +RPEs and -RPEs. For example, if increasing the size of +RPEs makes the FRN more positive, then increasing the size of -RPEs should not also make the FRN more positive: this is clearly necessary to prevent the lost value entailed by increasing -RPE size being confused with the gained value entailed by increased +RPE size.¹ This is equivalent to stating that there must be an interaction between RPE valence and RPE size, something we confirm in a meta-analysis of the FRN in Chapter 5.

While this interaction is a key criterion for identifying RPE encoding in the brain, it is the nature of the interaction that is of interest in determining whether processing of reward and punishment is integrated or dissociated. If the FRN represents the activity of an integrated RPE encoder, then the interaction should arise because voltage shifts negatively with increasing -RPE size and positively with increasing +RPE size. However, an interaction of RPE size and valence will also be observed if the FRN's voltage is responsive only to the size of *either* +RPEs or -RPEs. If the FRN were to

¹ The axiom is straightforward and we express it here in terms most suitable for the treatment in this chapter: that changes in the FRN's voltage in response to increasing RPE size should not be in the *same* polarity for both +RPEs and -RPEs. To prevent confusion however, we note that the axiom is expressed differently depending on how RPEs are treated. In Chapter 3, we used the metric of RPE *utility* rather than RPE size, whereby low -RPE size constituted high -RPE utility and vice versa. In this case, the axiom therefore states that that changes in the FRN's voltage in response to increasing RPE utility should not be in *different* polarities for both +RPEs and -RPEs. A further complication in the case of the likelihood-modulated FRN, is that likelihoods may refer either to the actual outcome (the convention used in this thesis) or to the likelihood of obtaining a desired outcome (e.g. Rutledge et al. 2010). Thus, when £5 can be won by rolling a six, and a six is not rolled, we refer to this outcome as a high likelihood -RPE, while Rutledge et al (who also use the axiomatic model) refer to this as a low likelihood -RPE. The consequence is that in that paper and those based on it, the axiom requires that changes in the FRN's voltage in response to increasing RPE utility should not be in *different* polarities for both +RPEs and -RPEs.

show such a response function this would provide further evidence of the dissociation of the processing of reward and punishment.

Unpacking the interaction of RPE size and valence is thus necessary to answer the question of whether reward and punishment are integrated or dissociated in the human brain. Normally such an unpacking would be achieved by examining simple effects, that is, by examining the extent of FRN sensitivity to RPE size for +RPEs and –RPEs separately. Many papers do report such simple effects, with a recent review of the literature performed by Walsh and Anderson (2012) reporting a significantly greater sensitivity of the FRN to changes in +RPE size compared to –RPE size, i.e. a preferential sensitivity to reward rather than punishment. Such an analysis of simple effects is nevertheless unsafe because the behaviour of single waveforms, on which simple effects are based, is subject to unknown component overlap. This can effect a positive or negative translation on only some of the single waveforms, distorting the pattern of simple effects. This problem was discussed in the introduction to Chapter 3 in terms of overlapping response functions, and in Figure 20 we show graphically how this might look in the case of an actual ERP. The figure shows an idealised FRN as an integrated encoder of both +RPEs and –RPEs, overlaid with a component that codes for an absolute, or unsigned prediction error. The summed effect of these components is a waveform which appears sensitive only to +RPE size. There is now mounting evidence of such salience encoding components in the same temporal interval as the FRN (Hauser et al., 2014; Talmi et al., 2013) and such components therefore stand to account entirely for the apparent preferential sensitivity of the FRN to +RPEs shown in Walsh and Anderson’s review.

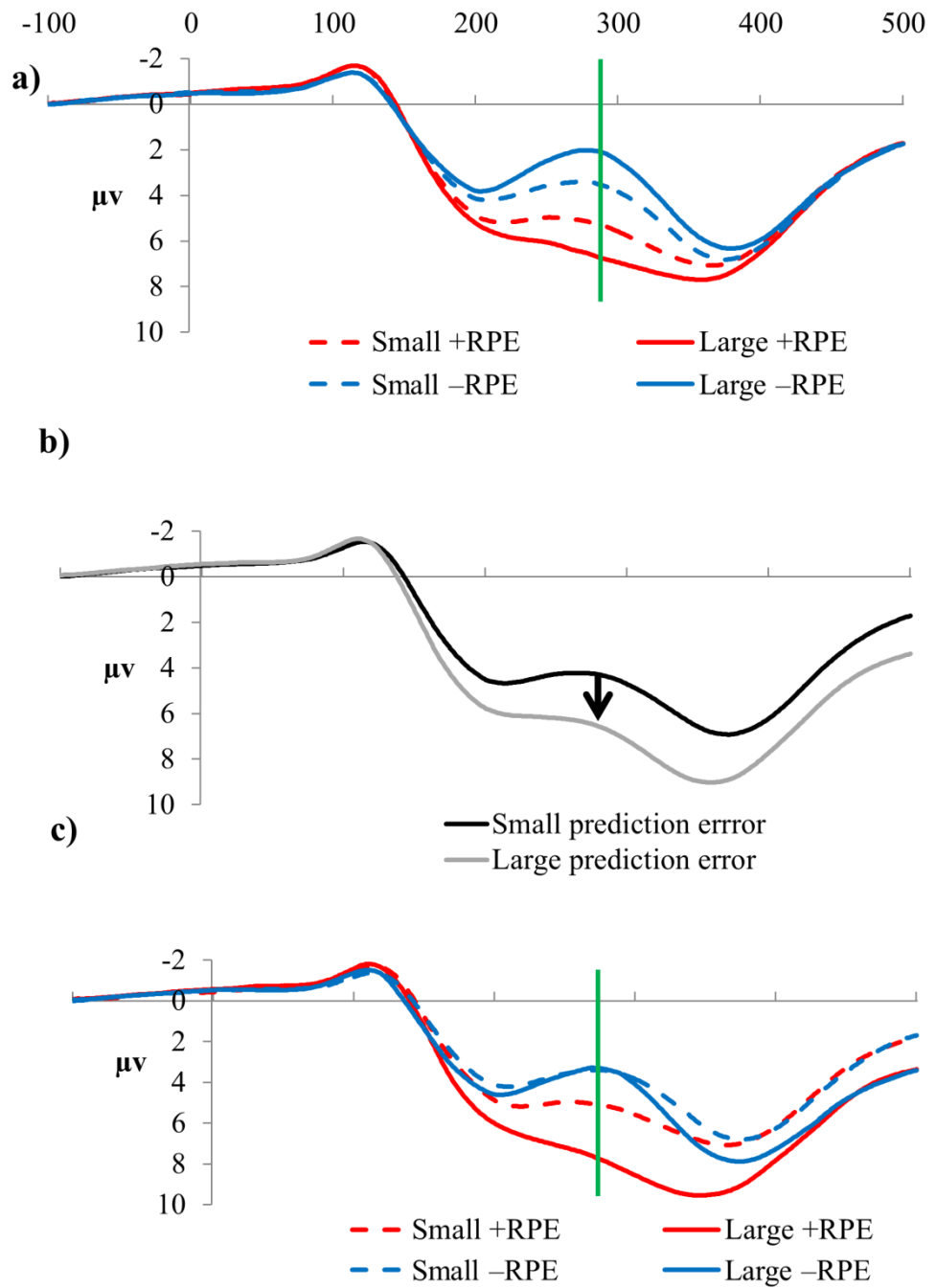


Figure 20 How component overlap undermines simple effects analysis of the FRN

How component overlap undermines simple effects analysis of the FRN (measured at the green line). a) How the FRN might look in simple waveforms unaffected by

component overlap. In this hypothetical case the FRN shows integrated coding, with both –RPEs and +RPEs modulated by their size, on a common utility scale where positive voltages are good. b) a separate component lies in the same interval which simply codes for the absolute size of the prediction error, i.e. its salience. Large prediction errors shift the waveform positively by ~2 μ v. c) The observed ERP waveforms. The effect of the salience encoding component is to shift large –RPEs and large +RPEs positively, producing equivalent voltages for small and large –RPEs while increasing the difference between the voltage of small and large +RPEs . Analysis of simple effects would result in the erroneous conclusion that the FRN responds to the size of +RPEs only.

In Chapter 3 we attempted to separate out the response functions by analysing activity across the whole waveform on a sample by sample basis, identifying transitions from one response function to another. We found that while integrated and salience response functions were both present, they had apparently not been combined into a spurious +RPE response function in the manner described above. In fact, while an integrated response was present insofar as polarity shifts were in the same direction for increased utility in both +RPEs and –RPEs (i.e. in opposite directions for increased RPE *size*, in keeping with the axiomatic model above) this overall effect was generated by consecutive responses to +RPEs and –RPEs individually, with salience responses occurring either side of these. Identifying this sequence of responses and their intervals would not have been possible using a single measurement interval, derived from the existing literature. Nevertheless, the merits of the technique used, and the robustness of its findings are still dependent on components being more consecutive than concurrent in time. That is to say that component overlap remains a potential problem, not least because its extent is unknown.

Because the true response function of the FRN will remain uncertain as long the component remains overlain by other components, a logical step would appear to be to separate out components using a technique such as principal components analysis (PCA) prior to examining simple effects. This technique was not used in earlier chapters since these used only a limited montage. Some previous attempts have been made in this regard. For example, Holroyd et al. (2008) concluded that the FRN responded only to +RPEs after spatial PCA revealed a frontocentral factor with this behaviour. However these authors did not follow the spatial step of the PCA with a further temporal step. Given surprisingly high variance explained by this factor (>50%), it is likely that this

would have allowed the separation of further frontocentral components. Foti et al. (2011b) did conduct a temporospatial PCA designed specifically to resolve the question of the relative sensitivity of the FRN to +RPEs and –RPEs. These authors found a factor whose temporospatial profile resembled the FRN, and which responded only to +RPEs. Crucially however, the experiment employed simple dichotomous good vs. bad feedback, and RPE size was not varied. This meant that the requirements of the axiomatic model of RPE encoding could not be satisfied, as this requires an interaction between RPE size and valence (Note: while it might be argued that failure to respond to –RPEs as well as +RPEs necessarily rules out the possibility that a component constitutes an RPE encoder, this is not the approach we take here and is not implied by the axiomatic model) .

This highlights a dichotomy within the literature between the FRN's operationalization and its proposed function. This component is characterised by a sensitivity to valence, 200–350 ms after feedback is received, which is maximal over frontocentral electrodes. These are the necessary and sufficient properties for its identification and, indeed, it is typically operationalised simply as a difference wave of good and bad outcomes. Sensitivity to RPE size does not constitute part of the operationalization. However, the theoretical claim of Holroyd and Coles that the FRN constitutes an RPE encoder *does* require observation of its response to RPE size. Thus Foti et al.'s (2011b) PCA study successfully isolated the FRN as typically defined, but nevertheless cannot be said to have successfully isolated an RPE encoding component. The distinction is far from trivial because in our meta-analysis in Chapter 5 we show that the interval of the feedback-locked ERP that shows an RPE size x valence interaction suggests the presence of multiple valence sensitive components, only some of which may meet the axiomatic requirements for an RPE encoder.

Thus, in this study our aim is to establish the axiomatic verities of the FRN by conducting a temporospatial PCA analysis of this component, similar to that of Foti et al.'s study, but including a factor of RPE size. To best implement this test, we parametrically varied RPE size across a wide range of values and established the correlation between voltage and RPE size on a sample by sample basis, as in Chapter 3. All the advantages of this approach described previously still pertain, however the approach is also particularly well suited to PCA. This is because when factorial designs are entered into a PCA, the relationship between conditions is not specified prior to extraction of components.. However, ignorance of the structure of the data can result in misallocation of variance during the PCA, especially in noisy data. Replacing a factor (here RPE size) with its *effect size* (here Pearson's r) ensures that this information is available to the PCA, ensures that all extracted factors explain variance in that factor, and no variance is "wasted" extracting factors that are merely obligatory responses to the arrival of a stimulus on a screen.

A number of other measures were taken to isolate the FRN and any other RPE encoding components as effectively as possible. Since RPE size is a function of both the magnitude of a reward or punishment and its prior likelihood, an RPE encoding component should be responsive to variations in either of these properties. Thus to ensure we were observing an RPE encoder rather than simply a response to likelihood or magnitude, we manipulated both these properties, rather than just manipulating magnitude as in the previous chapter. A second departure from the previous chapter was that rather than assigning the RPE value for a trial to simply the difference between an outcome's value and a single static EV for the block, RPE values were assigned with a modelling process that used a dynamically changing EV term which tracked recent interaction is considerably smaller than the interval showing a main effect of valence. reinforcement with a Rescorla-Wagner algorithm. While a case can be made for this

providing a more accurate assignment of RPE values than the procedure we used in Chapter 3, and indeed modelling of this kind is the norm in fMRI, the primary motive was methodological. This was that, under likelihood modulation, outcomes were a simple categorical variable of either reward or non-reward, and thus, even with the likelihood of reward being manipulated as 25%, 50% and 75%, RPEs were limited to three discrete values for each RPE valence, rather than being finely graded as they were under magnitude modulation. Modelling RPEs using a Rescorla-Wagner function allowed us to establish RPEs of a similar statistical distribution, thus allowing a fairer comparison of the effects of magnitude and likelihood modulation on the feedback-locked waveform.

Third, in this experiment, the task was modified to represent an instrumental conditioning task in which participants tracked the profitability of two possible responses (left and right key presses). This contrasts with the previous experiment in which participants were passive observers, told explicitly that actions had no effect on outcomes. The task was changed for two reasons. First, the grand average waveforms in that experiment, shown in Figure 15, showed a dramatic reduction in the overall FRN effect, and we feared that this would not survive further decomposition by PCA. Second, we perceive PCA to serve a remedial role in separating out components that have previously been described in agglomerated form, and we felt we could make the most impact by using an instrumental task since this was more representative of the FRN literature generally.

4.3 Methods

4.3.1 Participants

Eighty seven (23 male) students of the University of Plymouth participated for course credit and an opportunity to win money. All were right handed and under 29

years. Forty five participants were tested with RPEs manipulated by varying outcome magnitude, and forty two with RPEs manipulated by varying outcome likelihood.

4.3.2 Task

RPEs varied by magnitude. Apart from some modifications designed to make it an instrumental learning task, the task largely resembled that used in Chapter 3. It is summarized in Figure 21. Participants were shown a prompt and were required to press the right or left key of a key pad. A fixation cross appeared (600–700 ms duration), followed by the outcome stimulus (700 ms duration) and then a blank screen (800 ms duration). Participants performed the task in 32 blocks of 60 trials each.

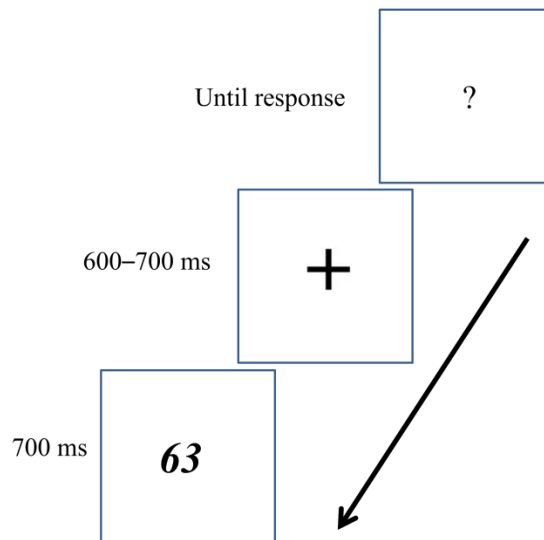


Figure 21 Summary of a single trial of the task procedure

Participants underwent gain domain blocks and loss domain blocks. In gain domain blocks participants were told that at the start of each block one of the keys was set to return outcomes with slightly higher values, but were not given the identity of this key. They were instructed to monitor the performance of the keys because their aim was

to accrue a sum of points of at least 3,000 in order to obtain a £2 reward for that block. In order to further motivate participants to monitor feedback, they were asked to estimate at the end of each block whether they had exceeded or fallen short of the 3000 target, and were awarded 20p each time their estimate matched the sum that was then reported to them. In fact, in every block each of the numbers 23 to 82 appeared once in a random order and the points total reported at the end of each block was a fictional sum set to be under or over 3,000 equally often. For loss domain blocks the task was identical except that participants were instructed to find the key returning lower numbers, since if the points total exceeded 3,000 they would lose £2.

Sixteen gain domain blocks were run followed by sixteen loss domain blocks, with this order reversed for half the participants. A blocked design was used for the domain variable since the study sought only to control, not study, domain effects and it was believed that alternating domain on a trial by trial basis would confuse some participants and reduce FRNs generally as feedback stimulus–reward associations were continually being reversed. Earnings for points accrued in blocks were a net zero for all participants. Earnings for the estimate of whether the 3,000 point target had been reached varied, with the average being £2.12

RPEs varied by likelihood. The task differed from the magnitude task in some key aspects required to realise the likelihood modulator, but in all other aspects as much similarity as possible was retained. Participants were told that key presses would always return either a one or a six: canonical values for this experiment. On each block, a guideline probability for sixes was given, which could take the value 25% (low), 50% (medium) or 75% (high). Participants were told that one of the keys had a slightly higher chance of returning sixes than the other, but were not given the identity of the key. On gain domain blocks participants were instructed to maximise the number of sixes obtained, since if they could reach a target number (set to 15, 30 or 45 for low,

medium and high likelihood blocks) they would earn £2. In order to further motivate participants to attend to feedback they were asked at the end of each block which key they believed had been set to have the greater likelihood of returning sixes, and were awarded 10p if this matched the key that was then reported to them.

In fact neither key was set to be preferable to the other and the experiment was designed to ensure blocks ended with the target number of sixes reached or missed equally often. Because it was possible that participants might actually be able to count sixes obtained, this number was reported faithfully. The procedure for the loss domain blocks was the same, with the exception that participants had to undershoot the target number of sixes to avoid losing £2. Thirty blocks were presented in total, with the domain reversed after 15 blocks. Within each domain, five sub-blocks were presented for each of the three baseline probabilities for sixes (i.e. low, medium and high). The order of these sub-blocks was counterbalanced. Earnings resulting from reaching the target number of sixes were a net zero for all participants. Earnings for stating which of the keys had been set to be better were an average of £2.20.

4.3.3 Modelling of RPEs.

While in the previous chapter RPEs were calculated with respect to a set EV of 52.5 here they were calculated as follows. At the beginning of each block, the expected value of the first trial was set to the mathematical expectation given the parameters for that block (always 52.5 under magnitude modulation and 2.25, 3.5 or 4.75 under likelihood modulation, depending on the baseline probability of sixes on that trial). Practice trials on the first block and experimental trials thereafter ensured that this was a reasonable estimate. Thereafter, EV was computed using a standard Rescorla–Wagner algorithm (Rescorla & Wagner, 1972). For each trial, t , an RPE $\delta(t)$ was computed as the difference between the actual outcome value $R(t)$ and its expected value $V(t)$ on that trial

$$\delta(t) = R(t) - V(t)$$

The expected value of the next trial $V(t+1)$ was updated by adding the RPE, weighted by a learning rate, α , set to 0.3.

$$V(t+1) = V(t) + \alpha \cdot \delta(t)$$

This value of the learning rate was arbitrary, but in keeping with similar studies of reinforcement learning (Metereau and Dreher, 2013; Seymour et al., 2005), and comparable with a value extracted in a recent study of non-reversal probabilistic learning and the FRN (Mas-Herrero and Marco-Pallarés, 2014). Reanalysis of the data using a learning rate of zero produced little change in the results.

4.3.4 EEG Recording

EEG data were collected from 61 Ag/AgCl active electrodes (actiCAP, Brain Products, Gilching, Germany) mounted on an elastic cap and arranged in a standard International 10-20 montage. Electrodes were referenced to the left mastoid and re-referenced off-line to the average of left and right mastoid activity. Vertical eye movement was monitored by a right suborbital electrode, and horizontal eye movement was monitored using an electrode on the right external canthus. Electrode impedances were kept below 20 k Ω . EEGs were amplified using a BrainAmp amplifier (Brain Products), continuously sampled at 500 Hz, and filtered offline with 60 and 50 Hz notch filters followed by a .1 Hz high pass filter and 30 Hz low pass filter. EEGs were time-locked to 100 ms before the onset of the feedback to 700 ms afterward, and then were baseline-corrected using the interval -100–0 ms. Eye movement artefacts were removed using a criterion of a voltage change exceeding 75 μ v/200 ms in eye electrodes. Other non-specific artefacts were removed using a criterion of any electrode showing either a voltage change exceeding 40 μ v/ms, a voltage value exceeding \pm 200 μ v relative to baseline, or activity across the epoch of <2.5 μ v. The percentage of trials removed was 85.8%. In case of electrode malfunction, where three or fewer electrodes were affected

in the course of an experiment, they were substituted using topographic interpolation (Perrin, Pernier, Bertrand, & Echallier, 1989). This produced an average incidence of topographic interpolation of 1.06%.

4.3.5 EEG and PCA Analysis

ERPs showing the categorical effect of valence, i.e. +RPEs vs. –RPEs, were created by conventional means, that is by averaging all trials of the given valence and then grand averaging those across participants, and are shown in Figure 22. However, statistical analyses were based on the parametric variation of RPEs, with separate correlation waveforms created for trials with –RPEs and +RPEs.

PCA was performed on these waveforms using the ERP PCA Toolkit Version 2.41 (Dien, 2010a), using identical means to those used by Foti et al (2011b) and following published guidelines (Dien, 2010b; Dien, Beal, & Berg, 2005; Dien, Khoe, & Mangun, 2007). Separate PCAs were performed for magnitude modulated and likelihood modulated RPEs, to ensure putative RPE encoding components such as the FRN were elicited in both cases, as theoretically required. First, a temporal PCA was performed using each time point as a variable and each combination of participant, electrode, and condition (+RPE vs. –RPE) as observations. Promax rotation was used. Temporal factors were discarded if they accounted for less variance than i) a factor extracted from a null, randomized dataset, i.e. if the factor failed a parallel test (Horn, 1965), or ii) a factor in the baseline, since these set an informal threshold for rejecting spurious factors. The extracted temporal factors were then subjected to spatial PCA using Infomax rotation. In this step, electrodes were used as variables, and each combination of participant, condition and temporal factor used as observations.

Following the method of Dien, Spencer, and Donchin (2003), factors were then reconstructed into waveforms using the product of the factor pattern matrix and the standard deviations. These could then be interpreted in the same manner as the

correlation waveforms depicted in Figure 23: indeed, waveforms in Figure 23 are simply the sum of the factors shown in Figure 24, minus the residual factors that were extracted in the PCA but not shown. The factor score for each participant x condition was taken as the amplitude at its peak value, following Foti et al. (2011b). These scores were used for t-tests on the sensitivity of the factor to +RPEs and –RPEs respectively.

It is important to note that because the waveforms entered into the PCA described sensitivity to RPE size, all extracted factors were also necessarily sensitive to RPE size. The response function of these factors was established by assessing their sensitivity to +RPEs and –RPEs respectively. For each of these two valences of RPE, one sample t-tests were performed across participants to establish whether the factor was sensitive to variations in RPE size for that RPE valence. The simple fact that the factor had been extracted indicated that sensitivity to RPE size was shown by at least one of the two valences, but not which. Significant sensitivity for +RPEs only would indicate a +RPE encoder, for –RPEs only a –RPE encoder. If the factor was sensitive to both RPE valences, but responded in opposite polarities for +RPEs and –RPEs it would be described as an integrated encoder. On the other hand, if a factor responded with the same polarity to both RPE valences, it could not be described as an RPE encoder at all, but rather a salience encoder.² In some cases, factors responded in the same polarity but with different strengths to +RPEs and –RPEs, typically stronger for the former. This differential sensitivity notwithstanding, such a factor could not be described as an RPE encoder because it could not distinguish between the two valences, producing an

² In the previous chapter we investigated response functions using the same rationale of separately testing sensitivity to +RPEs and –RPEs. However, in that study we correlated voltage against RPE utility rather than RPE size. As utility and size are inversely related in –RPEs, and positively related in +RPEs this meant that salience responses in that study were represented by *oppositely* signed +RPE and –RPE waveforms. In this study we have changed our convention to allow a direct comparison of our data with that of the study of Foti et al. (2011b) and the axiomatic model of Caplin and Dean (2008).

equivalent output for small +RPEs and large –RPEs, and thus failing the axiomatic model. Nevertheless, to assess the degree of this asymmetrical response, these salience encoding factors were additionally subjected to a paired-samples t-test to establish the degree of differential sensitivity to +RPEs and –RPEs.

After extraction of factors, source localisation was performed using Robert Oostenveld’s FieldTrip (<http://fieldtrip.fcdonders.nl/start>), which was interfaced directly from within the ERP PCA Toolkit. This used a four-shell model to specify a pair of hemispheric dipoles. To reduce the likelihood of finding a solution for a local minimum, a grid scan was performed on the head space to identify the best starting position for the dipoles. An iterative algorithm then moved dipole positions until maximum fit was achieved under a maximum-likelihood estimation algorithm. The entire epoch was used for the fitting process. Stability of the solution was assessed using the ERP PCA toolkit’s jack-knife technique, in which the spatial step of the PCA was recomputed as many times as there were participants, with a single participant omitted each time, allowing examination of the extent to which the solution depended on particular participants.

4.4 Results

4.4.1 Behavioural Results

To check whether participants engaged in the reinforcement learning task, the relationship between RPEs and subsequent switching behaviour was examined, as was done in Chapter 3. For all participants, RPEs were on average more negative on trials which were followed by a switch, with this effect significant for eighty-one of the eighty-seven participants. This suggests that participants attended to feedback and modified their behaviour accordingly. To allow a comparison with the data in Chapter 3, a group level analysis was performed concerning whether the RPE utility on a trial

affected the likelihood of switching to a different icon on the following trial. This showed that RPEs were significantly more negative on trials preceding a switch = (magnitude: $F_{1,101671} = 4950.51$, $p < .001$, $\eta^2 = .05$; likelihood: $F_{1,73538} = 1493.75$, $p < .001$, $\eta^2 = .02$) The effect sizes describing the degree to which participants adjusted their behaviour in response to prediction errors can be seen to be several orders of magnitude higher than in Chapter 3, justifying the claim that this a more instrumental task than that used in Chapter 3.

4.4.2 ERPs

Figure 22 shows the topography of valence coding using standard grand-average voltage waveforms, suggesting a typical frontocentral FRN at ~300 ms followed by a more parietal P3 response. Waveforms plotted at FCz show the relative negativity for bad outcomes that is characteristic of the FRN. It can be seen that the valence effect was much stronger than in the experiment described in the previous chapter.

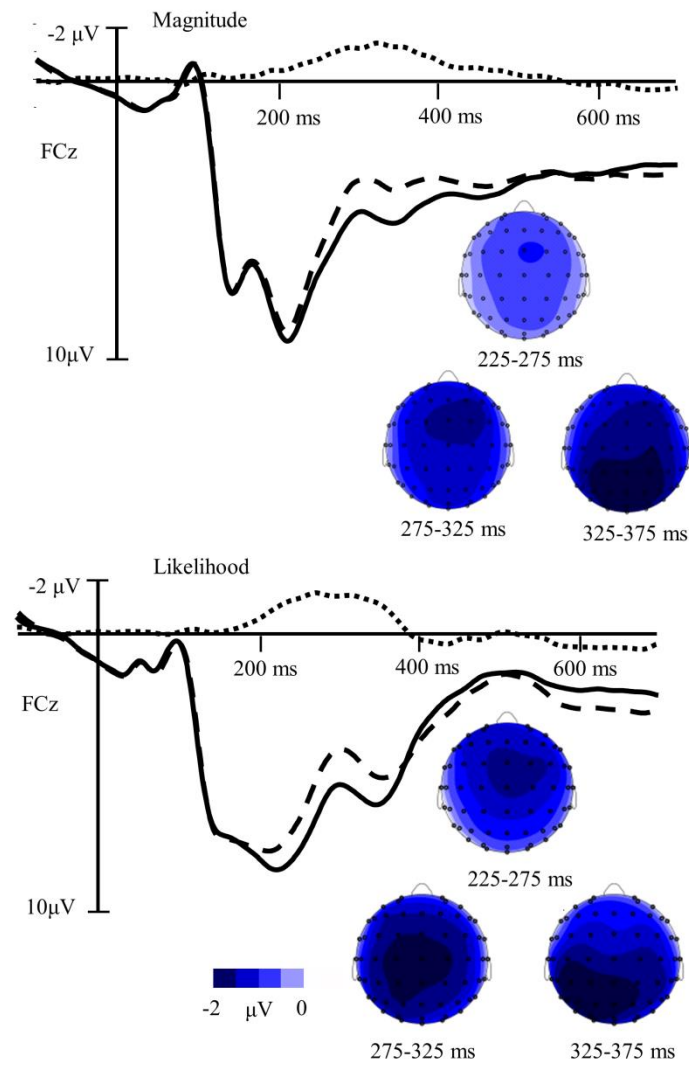


Figure 22 Grand average ERPs in response to +RPEs vs -RPEs

Figure 23 provides correlational plots showing the strength of encoding of –RPEs and +RPEs. Axes are plotted in reverse by convention. In contrast to Chapter 3, the y axis represents correlation of voltage with RPE size rather than RPE utility (plotting against utility would simply require reflecting the waveform for –RPEs across the x axis). For the majority of the epoch, the correlation between RPE size and voltage is positive for both +RPEs and –RPEs, indicating that increases in the size of both rewards and punishments produces a more positive feedback-locked waveform. This suggests that the underlying components are salience encoders, not RPE encoders. The exception is as ~330 ms where +RPEs retain a correlation with voltage but –RPEs show no correlation, suggesting that at this point an RPE encoding occurs, but for +RPEs only. This hypothesis was investigated in the PCA analysis below.

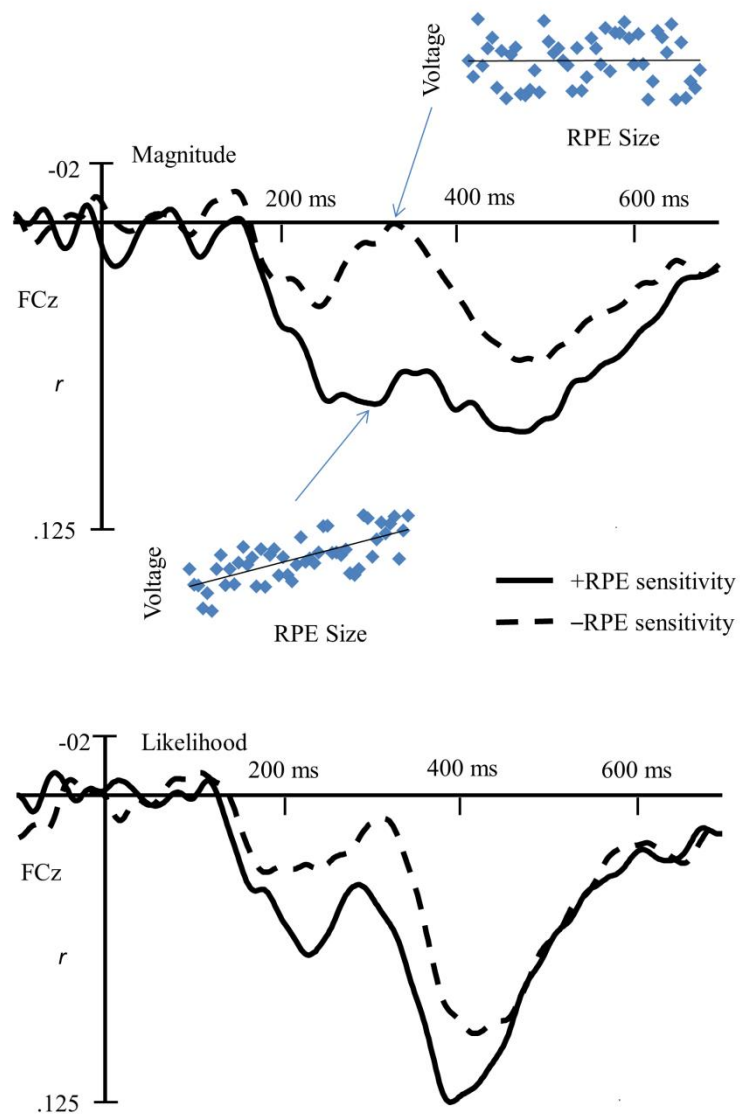


Figure 23 Correlational waveform prior to PCA

4.4.3 PCA Components

In the temporal step of the PCA, 11 temporal factors were extracted for magnitude modulated RPEs and nine for likelihood modulated RPEs. In the spatial step, three factors were extracted under both modulators producing a total of 33 and 27 temporospatial factors for magnitude and likelihood modulated RPEs respectively.

Four factors from each temporospatial PCA were selected for plotting and analysis on the basis that they accounted for three times more variance than the remaining factors. These factors are summarised in Table 2 and plotted for +RPEs and – RPEs separately in Figure 24. Of the remaining factors, the first temporospatial factor for each temporal factor is shown in Figure 25.

Table 2 Factor combinations selected for statistical analysis

Proposed Component	Modulator	Temporal loading peak	Spatial peak	Variance explained (%)	Response function	Factor
P2	<i>Magnitude</i>	256	Fz	6.89	Salience	TF2/SF1
	<i>Likelihood</i>	240	FC2	5.24	Salience	TF4/SF1
FRN	<i>Magnitude</i>	344	Cz	4.56	+RPE only	TF3/SF1
	<i>Likelihood</i>	322	Cz	6.09	+RPE only	TF3/SF1
P3	<i>Magnitude</i>	412	Cz	3.38	+RPE only	TF4/SF1
	<i>Likelihood</i>	418	Cz	16.99	Salience	TF2/SF1
Slow	<i>Magnitude</i>	532	POz	44.42	Salience	TF1/SF1
	<i>Likelihood</i>	528	P1	30.96	Salience	TF1/SF1

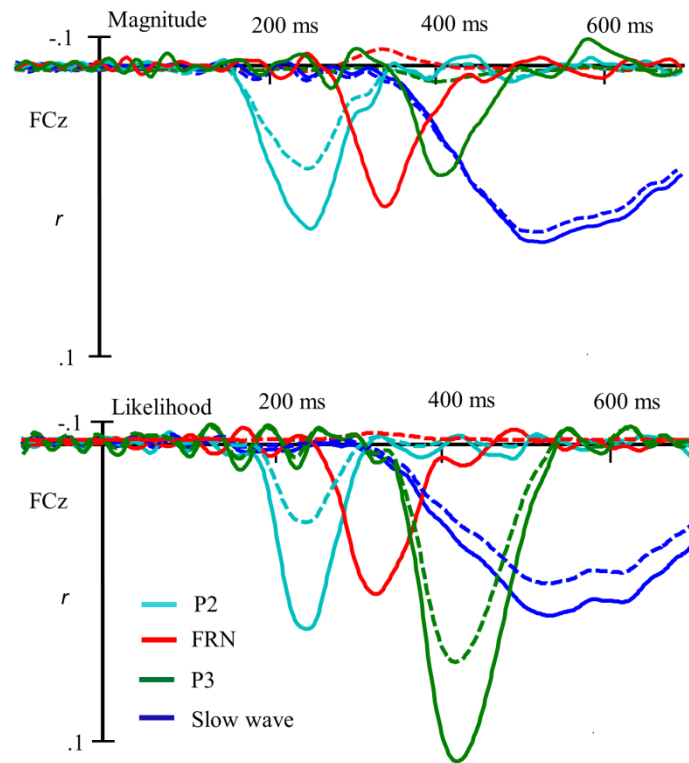


Figure 24 Factors extracted under PCA

Factors extracted using PCA. For each factor, dashed lines show the strength of its response to $-RPEs$, undashed to $+RPEs$.

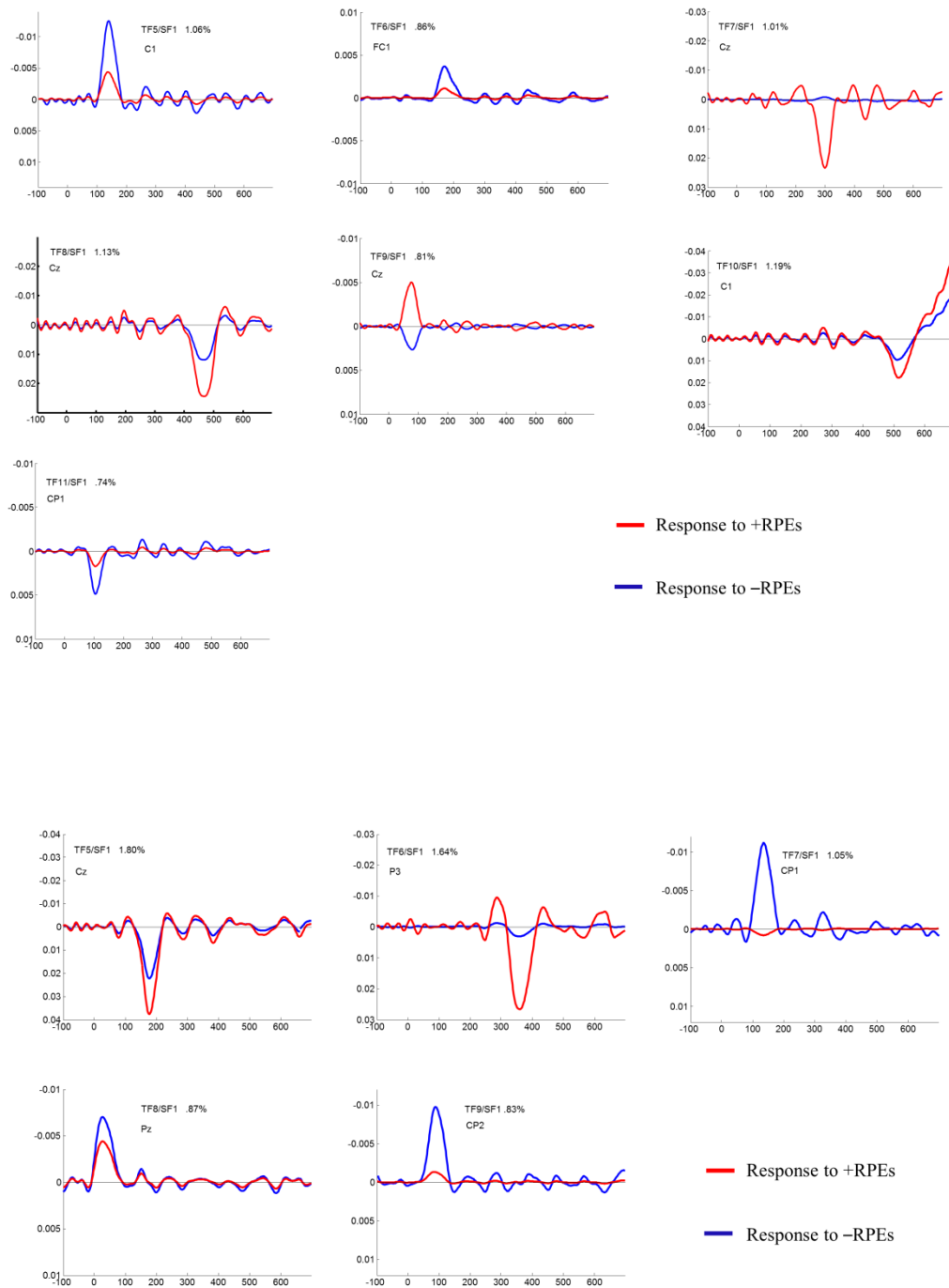


Figure 25 Factors extracted by the PCA but not selected for further analysis

Factors extracted by the PCA but not selected for further analysis, with the variance explained by that factor noted. a) magnitude, b) likelihood

Proposed FRN PCA component. Under both modulators, a factor was found that resembled the temporospatial profile of the FRN. One sample t-tests conducted at the peak of this factor's activity revealed it to be significantly responsive to +RPE size (magnitude $t(44) = 6.71$, $p < .001$; likelihood $t(44) = 7.97$, $p < .001$) but not to -RPE size (magnitude $t < 1$, likelihood $t < 1$). The response function of this factor, and its common occurrence in response to both magnitude and likelihood modulated RPEs was thus consistent with an RPE encoding function, albeit only of +RPEs. The slightly increased latency under magnitude modulation (344 ms vs. 322 ms) may be attributable to the greater complexity of processing a number in a sixty point range compared to the simple dichotomous stimuli of the likelihood experiments.

While a separate analysis of magnitude and likelihood modulated RPEs was necessary to establish their independent generation of the component, having established this, we collapsed likelihood and magnitude data together and re-extracted the factor to maximise the power of source localisation. This produced a good quality solution (residual variance = .98%) which identified a source in the putamen (Talaraich co-ordinates: 29.64, -17.87, -2.71). Application of the jack-knife technique showed this solution was extremely stable, with the standard deviation for these co-ordinates being 0.60, 0.54 and 0.42. The dipole was significantly larger in the right hemisphere $t(87) = 77.13$, $p < .001$. Figure 26 shows the factor and its source.

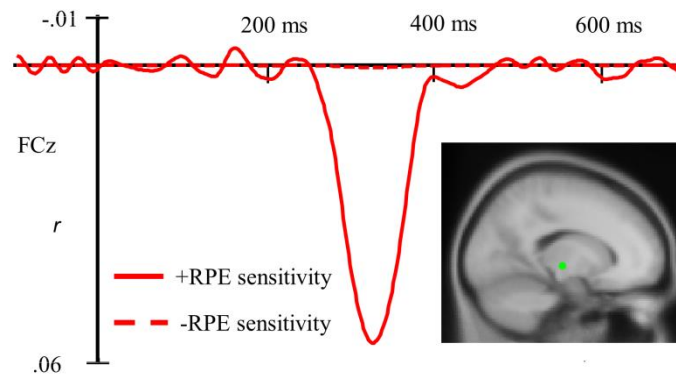


Figure 26 Factor TF3/SF1

Factor TF3/SF1, the proposed FRN factor, after collapsing together the magnitude and likelihood modulators, with the factor's estimated source shown

Other PCA components. A late parietal factor with a profile consistent with a slow wave, and showing a salience response function was found under both modulators. It was responsive to both +RPEs (magnitude $t(44) = 11.21$, $p < .001$; likelihood $t(41) = 8.74$, $p < .001$) and -RPEs (magnitude $t(44) = 7.99$, $p < .001$; likelihood $t(41) = 7.49$, $p < .001$), showing a greater response to +RPEs relative to -RPEs under the likelihood but not magnitude modulator (magnitude, $t(44) < 1$; likelihood $t(41) = 12.06$, $p < .001$).

A factor with a time course and amplitude consistent with the P3 was also found, though with a more anterior peak, at Cz. Under the magnitude modulator, the factor was responsive only to +RPEs (+RPEs, $t(44) = 7.66$, $p < .001$; -RPEs $t(44) < 1$). For likelihood, however, this component showed a salience response, having the same polarity of response to both +RPEs ($t(41) = 13.68$, $p < .001$) and -RPEs ($t(41) = 8.70$, $p < .001$), though with a greater response to +RPEs ($t(41) = 12.90$, $p < .001$). As such, its behaviour is not consistent with an RPE encoder.

In keeping with recent studies (Hauser et al., 2014; Talmi et al., 2013) a frontal salience encoding factor was found. Under the magnitude modulator, this had the same polarity of response to both +RPEs ($t(44) = 7.97$, $p < .001$) and -RPEs ($t(44) = 5.53$, $p < .001$) but with greater responsiveness to +RPEs ($t(44) = 3.43$, $p = .001$). The same pattern was observed under the likelihood modulator (+RPEs, $t(41) = 6.40$, $p < .001$; -RPEs, $t(41) = 3.06$, $p = .004$; +RPEs vs. -RPEs, $t = 12.37$, $p < .001$). Following Foti et al.'s (2011) classification of a similar factor shown under PCA, this has been provisionally dubbed the P2.

4.5 Discussion

The present study used principal components analysis to isolate components in the feedback-locked waveform, and so identify their individual responsiveness to +RPEs and –RPEs. A component with a temporospatial profile consistent with the FRN was extracted. Importantly, this component was not merely sensitive to the categorical valence of an outcome, i.e. whether it was better or worse than expected, but was sensitive to the actual size of the RPE, in the manner used by reinforcement learning models. Equally importantly, changes in voltage in response to RPE size were not in the same polarity for both +RPEs and –RPEs, behaviour that would violate an axiomatic model of RPE encoding. In fact, the component was responsive only to the size of +RPEs.

According to Holroyd & Coles' original theory, the FRN codes for a temporal difference error carried by dopaminergic midbrain neurons to the anterior cingulate cortex (ACC), where it serves as a teaching signal. In the present study, this component was responsive only to the size of +RPEs, suggesting that only a reward signal is transmitted. This finding is in keeping with the fact that midbrain dopaminergic neurons are strongly suspected to code +RPEs with scaled phasic bursts (Schultz et al., 1997), but because of their low tonic output appear to have little ability to code –RPEs with phasic reductions (Bayer and Glimcher, 2005). Other aspects of Holroyd and Coles' theory are less compatible with the present findings. In particular, the theory claims a close functional link between the FRN and the response-locked ERN, a component arising in response to internal registration of an error. Since the ERN necessarily arises exclusively from –RPEs, the plausibility of a common neural substrate with the FRN is strongly questioned.

The localisation of the FRN to the ACC is also not supported by the present data. While the ACC has been commonly implicated as the source of the FRN, this finding is

typically based on ERPs rather than factors extracted by PCA or a similar process. In contrast, recent source localisation attempts on PCA-extracted factors have consistently found a source in the striatum (Carlson et al., 2011; Foti et al., 2011b; Martin et al., 2009), as was found in our own study. This raises the possibility that source localisation attempts based upon FRN data that has not been decomposed by PCA may have been compromised by overlap with some other component. It is notable that the effect of valence in the undecomposed waveforms is strongest at 300 ms (shown by the difference waves in Figure 22) which is fully consistent with the literature, but that this effect nevertheless does not correspond to any extracted factor. Instead, this effect is a combination of TS3/SF1, the proposed FRN, and the earlier P2, which shows a salience response, but one which is somewhat stronger for +RPEs. Clearly, source localisation of the valence effect at 300 ms could thus not give a meaningful result. It is also worth noting that a source for the FRN in the striatum is consistent with fMRI studies showing a strong bias towards responsiveness to +RPEs in that structure. Interestingly, recent fMRI meta-analyses by Bartra et al. (2013), Garrison et al. (2013) and Liu et al. (2011) that specifically compared responsiveness to –RPEs and +RPEs, all found the striatum to be +RPE biased, but the latter two studies found the ACC to be more sensitive to –RPEs. Given that the FRN in the present study appears to be a +RPE encoder its localisation to the striatum rather than the ACC is consistent with this literature. Furthermore, these meta-analyses showed the areas of the striatum responding to +RPEs to be relatively much larger than other RPE encoding areas, increasing the plausibility of their producing a scalp detectable voltage, something which has been previously questioned (Cohen et al., 2011).

Our motive for using PCA in the present study was to remove components overlapping the FRN. An additional aim was to use this data-driven technique to reveal other components that appeared to encode RPEs. None were found. Instead, numerous

salience encoding components were found which appeared sensitive to absolute, or unsigned prediction error size. While some of these showed a larger response for +RPEs than -RPEs, thus satisfying the axiomatic model's first axiom, they failed its second axiom because voltage became more positive for *both* +RPEs and -RPEs as RPE size grew. The importance of the second axiom should be clear when one considers that the activity of these components will be the same for large -RPEs and small +RPEs, rendering them unsuitable as RPE encoders. The components had latencies, durations and strengths strongly resembling Foti et al.'s (2011b) PCA study, and also resembled the behaviour of ERP components shown in other recent studies and reviews (San Martin, 2012; Talmi et al., 2013).

One particularly notable absence in this study concerns the lack of any -RPE encoding component. Given the apparent value in coding the size of -RPEs, it is appropriate to consider the reasons for this absence. One possibility, also raised in Chapter 3, is that the extracted salience encoding components are in reality each a pair of separable components occurring at similar locations and latencies, which individually code for +RPEs and -RPEs, and in both cases do so by increasing their firing rate. Functionally, as long as the efferents of these components differ there would be no danger of confusing +RPEs and -RPEs. An advantage of having both components code increased RPE size by phasic bursts rather than having -RPEs encoded by dips, is that this allows for a more metabolically efficient low tonic firing rate for both RPE encoders. The EEG technique, with its relatively coarse spatial discrimination cannot definitively rule out such an explanation, though it becomes somewhat less plausible under PCA, where components are better separated. For the present argument, we simply note that it is less parsimonious an account of these components than the salience encoding account, since in each case it requires the invocation of two components rather than one and these must respond to RPE size in the same polarity.

There are, furthermore, prior reasons to expect salience encoding components to exist. They would appear to serve a clear alerting function (e.g. Bromberg-Martin, Matsumoto, & Hikosaka, 2010), and single cell studies, the only certain way to avoid confusing a real response function with one resulting from aggregated effects, have frequently found salience responses (Kobayashi et al., 2006; Matsumoto and Hikosaka, 2009).

The absence of any response to –RPEs in the present experiment contrasts with the experiment in Chapter 3, in which responses to both +RPEs and –RPEs were shown. The critical difference was that participants in this previous experiment were told that they could not influence the likelihood of reward. As such, none of the neural apparatus of instrumental conditioning should have been involved. The observed responsiveness to RPEs clearly implied that some expectation of value had been generated however. Following the classification of Balleine et al. (2008), it seems likely that these were Pavlovian values, predictive of the value of forthcoming reward, and updated by RPEs, but not associated with any behaviour by the participant. In contrast, in the present task, participants were told that outcomes were contingent on actions, and so instrumental values should have been generated, representing the value of an action (left or right key press). Of course, we would expect Pavlovian values to be generated in addition, however the absence of any –RPE Pavlovian value signal can be attributed to its being lost amid the much larger +RPE instrumental value signal (which was in this study seven times larger than the Pavlovian value signals in the previous experiment). Collectively, the pattern showed across our experiments, of +RPE and –RPE responses for Pavlovian values, but only +RPE responses for instrumental values, mirrors the findings in a recent behavioural modelling experiment (Huys et al., 2011) which showed just such a bias towards +RPEs in instrumental tasks. Other studies have also found that punishment is neglected during instrumental learning (Chase et al., 2011; Kahnt et al., 2009).

It should be acknowledged that, strictly speaking, the experiment in this chapter cannot be described as an instrumental task insofar as reinforcement was not in reality contingent on performance. However, it is a truism of psychology that it is the *beliefs* of a participant that have primacy in explaining behaviour or, as in the present case, neural activity. Indeed, the very large difference in effect sizes of the effect of reinforcement on behaviour in Chapter 3 and 4 (despite extremely similar tasks) bear testimony to the pre-eminence of beliefs in determining whether a task should be regarded as Pavlovian or instrumental when interpreting its findings

Why might there be a dearth of –RPE encoding for instrumental learning? One contributory factor in the present experiment, and others, may be that the task stressed deliberative action rather than inaction on the part of the participant. Boureau and Dayan (2011) observe that while, in principle, punishment and reward are orthogonal to action and inaction, in the real world animals typically act to obtain rewards and inhibit action to avoid punishments. There is widespread evidence that this has resulted in a heuristic coupling of reward incentivised behaviour with action, very possibly mediated by dopamine (Huys et al., 2011). The consequence is that in “go tasks” we expect +RPEs to be more readily generated than –RPEs, and in “no-go tasks”, the opposite to occur. Thus, the preponderance of go tasks in the literature may have placed undue prominence on +RPEs. However, another final possible reason for neglect of –RPEs might be that the size of +RPEs is simply more informative in the real ecologies in which reward circuitry has evolved. For example, in the case of food acquisition, animals will experience a finely graded degree of reward for each food item based on its quality. In contrast, all items that cannot be eaten are equally unrewarding. While a food patch or prey item can in principle be punishing, in that it gives a lower than expected return, the brute scarcity of food may mean that, in practice, all food acquisitions are generally +RPEs. In comparison, –RPEs in natural ecologies are most likely to be

threats. Since these are all relatively serious for long term fitness, a categorical response to a threat, with reduced sensitivity to graded size may be more appropriate. We concede that this conjectural but believe a consideration of the ecology of RPEs themselves is likely to inform the study of RPE encoders in the brain.

5 A Meta-Analysis of the FRN Using Great Grand Averages

The work in this chapter is strongly based on a published paper (Sambrook and Goslin, 2015)

5.1 Chapter Abstract

As an RPE encoder, the FRN should be sensitive to both the prior likelihood of reward and its magnitude on receipt. A number of studies have found the FRN to be insensitive to reward magnitude, thus questioning the Holroyd and Coles account. However, because of marked inconsistencies in how the FRN is measured, a meaningful synthesis of this evidence is highly problematic. We conducted a meta-analysis of the FRN's response to both reward magnitude and likelihood using a novel method in which published effect sizes were disregarded in favour of direct measurement of the published waveforms themselves, with these waveforms then averaged to produce "great-grand averages". Under this standardised measure, the meta-analysis revealed strong effects of magnitude and likelihood on the FRN, consistent with it encoding an RPE. In addition, it revealed strong main effects of reward magnitude and likelihood across much of the waveform, indicating sensitivity to unsigned prediction errors or "salience". The great grand average technique is proposed as a general method for meta-analysis of ERPs.

5.2 Introduction

5.2.1 A Meta-Analysis of the FRN

In Chapters 3 and 4, we presented evidence that the FRN met at least some of the requirements for an axiomatic model of RPEs of Caplin and Dean (2008). While absent from the model, an important further requirement of a *generalized* RPE encoder would appear to be that it respond to modulation of RPEs by both likelihood and magnitude. An early, influential paper by Yeung & Sanfey (2004) claimed that the FRN is in fact insensitive to variation in the magnitude of outcomes, and this claim has been echoed by a number of papers since. However, the claim contrasts strongly with our findings in Chapters 3 and 4. Rather than examining the merits and demerits of individual papers that have modulated the FRN with outcome magnitude, in this chapter we attempt to settle the matter through a meta-analysis. We also look at likelihood modulation of the FRN, partly for completeness, but partly to assess the evidence that a single component underlies both responses.

5.2.2 Existing Evidence for Modulation of the FRN by Magnitude and Likelihood

In their original paper outlining the RL-ERN, Holroyd and Coles (2002) confirmed that the FRN could be modulated by reward likelihood. While their claim that the FRN constituted an RPE has proven highly influential, at the time of its publication the supporting evidence was limited to this single experiment, with no examination of potential magnitude effects. Now, after more than a decade of research on the FRN, we are in a much better position to assess whether Holroyd and Coles' account is supported by the evidence.

Table 3 provides a summary of the evidence for magnitude and likelihood modulation of the FRN. While it is not an exhaustive review, since it only includes experiments that meet our criteria for the forthcoming meta-analysis, the table suggests that reward magnitude does not modulate the FRN in the predicted manner. For those studies manipulating magnitude, six experiments showed the expected effect, eight

studies reported no effect and three showed the opposite effect (i.e. the FRN was greater for low magnitude outcomes).

For those manipulating likelihood, the evidence is stronger, if still not entirely consistent, with thirteen studies showing the predicted effect and six reporting no effect.

A similar review by Walsh and Anderson (2012) mirrors this picture.

Concerning the likelihood modulator, a simple sign test applied to 25 studies showed a significant effect consistent with an RPE coding. In comparison, magnitude could not be shown to significantly modulate the FRN. The authors argued the absence of magnitude effects could be because the majority of experiments cued participants as to whether an outcome would be high or low magnitude at the beginning of each trial. Thus, magnitude effects in these experiments could have been lost to scaling effects. The two studies in Walsh and Anderson's review that were un-cued showed at least partial support for an FRN modulated by magnitude, and on this basis the authors argued support for the Holroyd and Coles theory. However this very limited sample must be acknowledged to leave any meta-analytical basis for the magnitude modulator unproven.

5.2.3 Problems with FRN Measurement and Implications for Meta-analysis

The ERP technique's poor spatial resolution means that any individual experiment using ERPs is vulnerable to spurious conclusions arising from overlap of the component under consideration (here, the FRN) with other components occurring in the same temporal interval. The difference wave methodology by which the FRN is best studied will not remove interfering components which also code valence, nor can it remove all the effects of components that are even partially affected by valence. In gathering the data for this meta-analysis, examination of individual studies' waveforms showed a remarkable variability in their character, assumed to arise from differences in task, procedure and stimuli. This suggests that the FRN suffers from serious overlap

with unknown components which might, quite incidentally, be responsive to any of the three factors (valence, magnitude, likelihood) under study. Because the sources of the component overlap are unknown in each case, a broad meta-analysis is therefore more robust than any single experiment.

A serious hindrance to meta-analysis however, is the lack of consistency in how the FRN is quantified. In some papers it is measured by the voltage of a single peak, in others the difference between two peaks, in others the difference of one peak and the average of surrounding peaks, and in others by the mean amplitude in a set interval. Analysis is sometimes conducted on difference waves and sometimes on the original simple waveforms. Perhaps most seriously, the temporal interval in which the FRN is measured varies widely. Table 3 shows the quantification of the FRN in each of the studies used in this meta-analysis. It was found that both mean amplitude measures and peak assignment are made in intervals ranging from 50 to 150 ms duration, at substantially different latencies, with some studies using intervals that do not even overlap. The interval 200–350 ms after feedback, where the bulk of the FRN measures lie, is characterized by a steep, alternating, positive-negative-positive going waveform, and so differences in the interval in which the FRN is measured can have large ramifications, with waveforms from different experiments that look similar on the page producing opposing conclusions once they are quantified and subjected to a statistical test. This may lead to failures of replication. Conversely, it is also possible that unexpected effects may go unnoticed as a result of the interval used, leading to successful but unwarranted replication, and an inflating of the apparent robustness of the FRN component and effects associated with it. The consequence for meta-analysis is that the compilation of the *statistical results* of FRN studies would be far more sensitive to idiosyncrasies in how the FRN is quantified than is desirable.

5.2.4 A Novel Method: Great Grand Averages

The problem of miscellaneous measures described above prompted us to use a novel means of meta-analysis. Typically, a meta-analysis uses standardized effect sizes derived from individual research articles as replicates for some statistical test for an effect of interest (most typically whether the effect size is significantly different from zero). A basic assumption of any statistical test, meta-analytic or otherwise, is that all data constitute observations of the same phenomenon. In our case this phenomenon would consist of neural events comprising activity of the FRN component. As previous discussed, it is not well established whether the neural events measured across the range of previous studies arise exclusively from the FRN. Moreover, the mixture of mean voltage and peak to peak measures and the wide range of intervals used to quantify the FRN component also raise serious concerns over the equivalence of effect sizes measured across the literature.

Our meta-analysis avoided the miscellaneous measures problem by ignoring the quantification of the FRN the authors of individual papers had used. Instead, we took published grand average waveforms in all experiments that met our criteria, digitized these waveforms to extract their co-ordinates and averaged these coordinates across experiments to create composite waveforms representing “great grand average” (GGA) waveforms. While this approach is borne of necessity, there are nevertheless some benefits to bypassing the quantification provided by the original authors and going “upstream” to the published waveforms. One advantage is that a great deal of information is thrown away in the conversion of waveforms to unitary scores of component amplitude, and the GGA technique retains that information up to the point of quantifying the final GGA waveform. This means that effects that are small or lie outside typically analyzed intervals, but which are consistently present, can become noticeable.

A second advantage is the method's potential to reduce the effects of component overlap. In a single ERP experiment, averaging across trials reduces the effect of incidental neural activity that is peculiar to a given trial, thereby accentuating task-related components, which are elicited on every trial. However, this does not help reduce components that happen to be elicited by the task, but are not the subject of the experiment. This causes component overlap, and complicates the measurement of the component under study. Under the GGA technique some of this component overlap is reduced due to the variations in the tasks used in different experiments. Averaging across experiments reduces the effect of incidental components which are peculiar to a given task, thereby accentuating the component under study, which will be elicited on every trial. Of course, components other than the one under study which happen to be elicited by the *factors* of an experiment (here valence, magnitude and likelihood) will not be reduced by the GGA technique.

A third advantage arises from differences in the latency at which the FRN occurs in the pool of experiments used to create the GGA waveforms. This produces “smearing” of the peaks that characterize the feedback-locked ERP, reducing their amplitude and widening the duration of general positive and negative deflections. For the purposes of the study at hand, we regard such smearing as a methodological strength. This is because it reduces the availability of bespoke intervals in which strong, but likely unreplicable effects can be found. This ensures a fairer test of the RPE account. Thus what is lost in (possibly misleading) peak amplitude is gained in reliability. It must be noted that the advantages of this meta-analysis technique are not specific to the study of the FRN, rather, they are highly generic and could be applied to the meta-analysis of any ERP component.

The disadvantages of the method are that the extraction of the data directly from the waveforms introduces a new source of measurement error, and that disregarding the

reported statistics eliminates the only source of information concerning the within-study variance of the studies entering the meta-analysis. These issues are considered empirically later.

5.2.5 Moderators

While the variability of the waveforms produced by FRN experiments has complicated their interpretation and presented methodological challenges, it is possible that some of this variation is systematically related to differences in experimental tasks, and can thus be used to infer properties of the component. We therefore performed the following moderator analyses.

RPE modulator. Modulator refers to whether outcome magnitude or outcome likelihood was used to manipulate the size of RPEs. While demonstrating that the FRN is a generalized RPE encoder requires that it be responsive to both modulators, and so their effects needed to be established independently, a comparison of those effects is potentially illuminating since evidence that the FRN is a generalized RPE encoder would be bolstered by a relative insensitivity to the source of the RPE size modulation.

Control over outcome. The expected value against which an RPE is generated might consist either in the expected value of the preceding stimulus, or in the expected value of the action performed in response to that stimulus (Balleine et al., 2008). That is to say, the RPE might contribute to either Pavlovian or instrumental conditioning. This matter may be addressed by examining the degree to which control over outcome affects FRN amplitude. An RPE used in Pavlovian conditioning, perhaps reflecting a general role in valuation, will occur even when participants passively observe outcomes. In contrast if the FRN is greatest following a meaningful action on the part of the participant this suggests a role in instrumental conditioning.

Magnitude cueing. While reward likelihood is fundamentally limited to values between zero and one, there is no such delimitation to reward magnitude. In order to be

able to show satisfactory discrimination across a wide range of outcome magnitudes it would appear necessary that the FRN scale its response relative to the range of magnitudes considered available in the immediate context. Such scaling has been shown by Padoa-Schioppa and Assad (2006) and by Tobler, Fiorillo, and Schultz (2005) in macaque midbrain neurons, which produced equivalent responses to rewards of different magnitudes when the range of magnitude available on that trial was signaled to the participant beforehand. Bunzeck, Dayan, Dolan, and Duzel (2010) found a similar scaling effect in a study of humans using fMRI. Schultz (2009) has suggested that scaling is performed not on the absolute range of outcomes possible in a given context but on their estimated distribution, and that RPEs accordingly represent z scores. While the question of scaling is of theoretical interest, it is also methodologically important since scaling effects may have been responsible for the absence of FRN sensitivity to magnitude in the literature. This moderator analysis investigated whether effects of magnitude on the FRN were reduced in cued experiments as would be expected if scaling occurred.

Domain. While the FRN has a well-established sensitivity to valence, that is the sign of an RPE, this is formally orthogonal to the domain of the outcome, that is, whether the outcome constitutes an actual monetary loss or gain. For example, losses can still be +RPEs if they are smaller than expected losses. In Chapter 2 we showed a relative sensitivity of the FRN to the gain domain when domain was highlighted by the task, but that this effect was absent when domain was rendered less apparent by blocking the domain variable in Chapter 3. The possibility of a neural dissociation of how outcomes are processed in gain and loss domains is of broad theoretical interest, not least because this reduced sensitivity for outcomes in the loss domain, or “loss indifference”, is in direct opposition to the prediction of loss aversion made by prospect theory.

Table 3 Experiments used in the meta-analysis

Experiments used in the meta-analysis, including whether an effect consistent with an RPE-FRN was found, FRN quantification, the waveform used if additional ones were present in the listed figure (WAV), the domain of the gamble (DOM) and whether magnitude of an outcome was cued (CUE)

Experiment	N	RPE-FRN ^a	Site	FRN ^b	WAV	FIG	DOM	CUE ^c
<i>Likelihood</i>								
<i>Modulator</i>								
Bellebaum & Daum (2008)	17	Yes	Pool	Mean amp 220–280	Post-learning	4	Gain	
Bellebaum et al. (2010a)	15	No	Pool	Peak to peak: P2 (150–N2) to N2 (200–340)	Post-learning	4b	Gain	
Bellebaum et al. (2011)	18	Yes	Cz	Peak of diff wave 100–300	Younger participants	3a	Gain	
Franken et al. (2010)	47	–	Fz	Mean amp 200–300		2	Gain	
Hajcak et al. (2005) Expt 1	17	No	Fz	Peak of diff wave 200–500		1	Gain	
Hajcak et al. (2005) Expt 2	12	No	Fz	Peak of diff wave 200–500		3	Gain	
Hajihosseini et al. (2012)	20		Fz	Mean amp 100 either side of N2 peak		2a	Mixed	
Holroyd et al. (2003)	10	Yes	FCz	Peak to peak: P2 (160–240) to N2 (P2–325)		2c	Gain	
Holroyd et al. (2008)	12	–	FCz	Peak to peak: P2 (160–240) to N2 (P2–325)	Time-estimation	1b	Gain	
Holroyd et al. (2009) Expt 1	20	Yes	FCz	Peak of diff wave 0–600		2	Gain	

Holroyd et al. (2009) Expt 2	15	Yes	FCz	Peak of diff wave 0–600		2	Gain
Holroyd et al. (2009) Expt 3	15	Yes	FCz	Peak of diff wave 0–600		2	Gain
Holroyd et al. (2011)	18	Yes	FCz	Peak of diff wave 200–300	Outcome locked	1a	Gain
Larson et al. (2007)	11	–	FCz	Peak to peak: P2 (125–325) to N2 (P2–325)	Control	2	Gain
Liao et al. (2011)	19	Yes	Fz	Peak of diff wave 150–500	Outcome locked	4	Gain
Morris et al. (2011)	23	No	Cz	Peak of diff wave ~180–300	Passive gambling task	1	Gain
Pfabigan et al. (2012)	20	No	FCz	Peak to peak: P2 (preceding positive peak) to N2 (200–350)	Second half	1	Mixed
Sambrook & Goslin Chapter 4	42		FCz	Unpublished			Gain/ Loss
Talmi et al. (2013)	20	No (p<.06)	Pool	Amplitudes at sample points 205–250	Reward condition	4 ^d	Gain
Walsh & Anderson (2011a)	13	Yes	FCz	Mean amp of diff wave 200–300		4	Gain
Walsh & Anderson (2011b)	20	Yes	FCz	Mean amp of diff wave 200–350	No instruction condition	3	Gain
Walsh & Anderson (2013) Expt 1	14	Yes	FCz	Mean amp of diff wave 240–400	Standard and novel	6	Gain
Walsh & Anderson (2013) Expt 2	14	Yes	FCz	Mean amp of diff wave 240–400	Standard and novel	6	Gain
Walsh & Anderson unpublished Expt 1	13		FCz	Unpublished			Gain
Walsh & Anderson unpublished Expt 2	13		FCz	Unpublished			Gain
Yu et al. (2011)	16	–	Fz	Mean amp 275–325 and peak to peak (details	Outcome	2b,c	Mixed

not given)
locked 25%
/ 75%

Magnitude Modulator

Banis & Lorist (2012)	32	Wrong way	FCz	Mean amp 230–300 / Mean amp 230–300 relative to average of mean amps of P2 (180–225) and P3 (320–390) / Peak to peak P2 (150–230) to N2 (P2–330)	Average of noise	2	Mixed	N
Bellebaum et al. (2010b)	20	Yes	Fz	Peak to peak: P2 (preceding positivity from 150) to N2 (200–350)	Blocks 3–6, 5c vs. 50c	3b	Gain	Y
Gu et al. (2011)	24	Yes	Fz	Peak to average peak: P2 (preceding positive peak) to N2 (200–400) to P3 (succeeding positive peak)	Outcome valence subsequently	4b	Mixed	Y
Hajcak et al. (2006) Expt 1	16	No	Fz	Peak to peak: P2 (150–350) to N2 (P2–350)		1	Mixed	N
Hajcak et al. (2006) Expt 2	17	No	Fz	Peak to peak: P2 (150–350) to (P2–350)		3	Mixed	N
Hajihosseini et al. (2012)	20	No	Fz	Mean amp 100 either side of N2 peak		2c	Mixed	Y
Kamarajan et al. (2009)	48	Wrong way	FCz	N2 peak 200–275	Sexes averaged	4	Mixed	Y
Luo & Qu (2013)	18	Yes	FCz	Mean amp 200–250	Win/loss at ¥1 vs. ¥40	3a,b	Mixed	Y
Mushtaq et al. unpublished	29		FCz	Unpublished			Mixed	Y
Nittono et al. (2008)	16	–	Fz	Peak to peak: P2 (preceding positive peak) to N2 (150–300)	Even cond., –10/–1/+1/+10	1a	Mixed	N
Onoda et al. (2010)	17	–	FCz	Peak of diff wave 250–400		1	Mixed	Y
Pedroni et al. (2011)	16	–	Cz	TANCOVA over entire waveform		From author	Gain	Y

Roberts et al. (unpublished)	26		Fz	Unpublished			Mixed	N
Sambrook & Goslin (2014)	55	Yes	Pool	Correlation of voltage and utility over entire waveform		From author	Gain/ Loss	N
Sambrook & Goslin unpublished Expt 1	48		FCz	Unpublished			Gain	Both
Sambrook & Goslin Chapter 4	45		FCz	Unpublished			Gain/ Loss	N
San Martin et al. (2010)	22	–	FCz	Mean amp 240–310		4	Mixed	Y
Santesso et al. (2011)	30	Wrong Way	FCz	N2 Peak (200–400)	Adult participants	2	Mixed	Y
Sato et al. (2005)	18	No	Fz	Peak to peak: P2 (150–220) to N2 (P2–325)		1	Mixed	Y
Schuermann et al. (2011)	20	Yes	FCz	Peak to peak: P2 (100–300) to N2 (200–400)		2	Mixed	Y
Talmi et al. (2013)	20	–	Pool	Amps at sample points 205–250	Reward	4 ^d	Gain	Y
Toyomaki et al. (2005)	13	–	Fz	Peak to peak: P2 (unspecified) to N2 (unspecified)	–500/–10/ +10/+500	2	Mixed	Y
Van den Berg et al. (2011)	42	–	Fz	Peak to peak: P2 (150 –350) to N2 (following negative peak)		2	Mixed	N
Wu & Zhou (2009)	16	No	FCz	Mean amp 250–350	Expected magnitudes	1	Mixed	Y
Yeung & Sanfey (2004)	16	–	FCz	Peak to average Peak: P2 (preceding positive peak) to N2 (200–400) to P3 (succeeding positive peak)		2	Mixed	Y
Yi et al. (2012)	28	No	Fz	Peak of N2 (200–400)		4b	Mixed	Y
Yu & Zhou (2006)	20	No	Fz	Mean amp 25 before and after peak of diff wave	Execution	1	Mixed	Y
Yu & Zhou (2009)	14	No	Fz	Mean amplitude 200–300	"To bet" trials	3	Mixed	Y
Zottoli & Grose-Fifer (2012)	18	Yes	FCz	Peak to peak: P2 (150–300) to N2 (200–425)	Adult	2a	Mixed	Y

^a A dash indicates that the RPE-FRN was not reported.

^b Values in parentheses indicate the interval in which peak assignment was made in milliseconds.

^c In cued studies participants know the magnitude of the forthcoming feedback but not its valence, in un-cued studies participants knew neither its magnitude nor its valence.

^d Eight waveforms corresponding to the valence x magnitude x likelihood design were given: these were all digitised and the unwanted factor collapsed out by averaging pairs of waveforms

5.3 Methods

5.3.1 Inclusion and Exclusion Criteria

The independent variables used in this meta-analysis were outcome valence (+RPE, -RPE), outcome magnitude (high, low) and outcome likelihood (likely, unlikely), where this final variable refers to prior likelihood of an obtained outcome. For inclusion, an experiment had to contain within it a 2 x 2 factorial manipulation of valence with respect to either likelihood or magnitude. Where more than two levels of the likelihood or magnitude variable were presented in a paper, intermediate ones were ignored in order to maximize contrasts.

The dependent variable differed depending on the particular contrast examined, as detailed in the coding procedures section below. In all cases it was a voltage derived from the differencing of four simple waveforms related to the factorial design described above. Consequently, a key inclusion criterion was that a study must present such a set of simple waveforms. Waveforms had to be plotted for at least 500 ms post-feedback and 100 ms prior, and had to be locked to feedback, not response. Since waveforms were in many cases plotted at only a single electrode, and since variability of the electrode used suggested a broad distribution for the FRN, an experiment was included as long as it presented waveforms at Fz, FCz, Cz or “a frontocentral pool”. Variability was minimized by using FCz waveforms where available (even if individual papers reported the FRN to be maximal at a different site), and where they were not, using Fz, Cz or frontocentral pool in that order of preference.

Studies using populations other than healthy non-older adults were used only if control data for this population were available and participants had not been selected on the basis of any pre-screening (e.g. personality scales). The experiment had to offer monetary rewards conveyed by feedback, though tasks could vary widely, including guessing games, time estimation tasks and simply passive observation. Experiments

could either employ mixed gambles comprised of wins and losses, gain domain gambles where participants either won or failed to win a stake, or loss domain gambles where participants lost a stake or successfully avoided this. Losses and omitted rewards were classed as -RPEs, wins and avoided losses were classed as +RPEs. Where separate waveforms were presented for the portion of an experiment before and after participants learned a rule that allowed them to assess reward likelihood, waveforms for the portion after learning were used, since these could be expected to produce the strongest effect of likelihood on prediction errors. Experiments which manipulated factors other than the three of interest were included, though in some cases waveforms were used at one level of that additional factor, often a control level, if available and appropriate (see Table 3).

Experiments were excluded if the factor of magnitude, likelihood or valence was confounded with another variable. While experiments manipulating both magnitude and likelihood were acceptable (and in these cases were used twice in the analysis, once for each modulator) they were excluded if these variables confounded each other. This was common in Iowa Gambling Tasks and where participants could genuinely optimize their choice. Experiments where the FRN was a response to observation of another's performance were excluded. Magnitude experiments were considered ineligible if levels of the magnitude variable were blocked, since we expected this would strongly exacerbate scaling effects, with the FRN responding simply to the valence of the outcome at the given level of the stakes in that block (though in fact no otherwise eligible experiments were excluded on this basis). In the case of likelihood experiments, the following criteria were employed. There had to be two levels of the likelihood modulator either side of, and equal distance from 50% probability to avoid confounding likelihood with uncertainty, a property the feedback-locked ERP may be responsive to (Yu, Zhou, & Zhou, 2011). If participants received explicit instruction on probabilities this had to be consistent with real probabilities so that there could be no ambiguity

regarding the value of expected value that RPEs were generated with respect to.

Experiments were excluded if participants could learn to actively avoid disadvantageous trials (e.g. by knowing which button to press to always ensure >50% reward probability) since this made unlikely +RPEs and likely –RPEs infrequent, introducing a possible confound, and also leaving participants’ motives for their sub-optimal choice unclear.

5.3.2 Moderator Analyses

Modulator. Coding of this moderator was straightforward. In those cases ($k = 2$) where both magnitude and likelihood were manipulated, the two conditions were entered as independent studies for this analysis. Sample sizes were $k = 27$ for magnitude, $k = 24$ for likelihood.

Control over outcome. Operationalizing this moderator was inherently problematic because perception of control is highly variable across people (Langer, 1975). We used three levels of this moderator variable, which we believed would maximize contrasts. Level 1, termed “passive”, covered tasks in which participants were given no opportunity to act meaningfully prior to feedback. At the other end of the scale, Level 3, termed “rule implementation” comprised tasks where actions could be performed and where feedback was genuinely (and therefore ultimately visibly) dependent on choice of action. Level 2 was termed “guessing”. This level encompassed all tasks in which participants acted but could not actually affect the outcome. This included, for example cases where participants had to guess the location of a prize, or choose the stakes for a particular trial. It is true that participants might have *believed* that they had a degree of control over the outcome, but information on these beliefs was generally not available. We assumed that where control over an outcome was neither evidently absent (Level 1) nor present (Level 3), participants would experience, on average, some intermediate perception of control. Sample sizes were $k = 5$ for Level 1, $k = 24$ for Level 2, $k = 12$ for Level 3.

Domain. A direct comparison of the amplitude of the FRN in loss and gain domains could not be made since only two studies included any pure loss domain trials. However many studies offered mixed gambles, in which +RPEs were always gains and –RPEs always losses, and the loss indifference effect described earlier might be expected to attenuate effects in the loss portion, producing a net reduction of the FRN overall in mixed gambles. Domain was therefore coded with two levels. The first, “gain domain” comprised all cases where the worst possible outcome was no reward. The other level, “mixed domain”, comprised cases where monetary losses as well as gains could be incurred. Sample sizes were $k = 25$ for gain, $k = 26$ for mixed

Magnitude cuing. This analysis applied to magnitude studies only. Cued studies comprised all cases where participants knew the magnitude of the forthcoming feedback but not its valence, un-cued studies comprised cases where they knew neither its magnitude nor valence. A single study in which magnitude cuing was manipulated as an independent variable was entered into this analysis as two separate studies. Sample sizes were $k = 20$ for cued, $k = 8$ for uncued

5.3.3 Search Strategies

Published data. A search for English language journal articles and books was performed using the following databases: PsychInfo, PsychBooks, PsychArticles, ERIC, PubMed and Web of Science. Results were compiled in EndNote. Abstracts, titles and keywords were searched using the term "feedback negativity" OR "feedback related negativity" OR "feedback error-related negativity" OR "reward positivity" OR "feedback correct related positivity". Duplicates, clearly inappropriate journals and conference abstracts were removed without inspection, as were papers published prior to 1997 (the year of publication of the first FRN paper, Miltner et al., 1997). Two hundred and fifteen papers remained, of which 42 were deemed eligible after checking inclusion and exclusion criteria.

The FRN is sometimes referred to generically as an “error related negativity”, even though this term is more commonly used to refer to a waveform locked to participants’ own responses, and indicating internal registration of a known error, rather than a response to external feedback. It is also sometimes referred to generically as a mediofrontal negativity. We conducted a secondary search using the term “error related negativity” OR “mediofrontal negativity” OR “medial frontal negativity”. After removing duplicates, duplicates with the earlier search, clearly inappropriate journals and conference proceedings, and papers predating 1997, 1012 papers remained. The abstracts were scanned for evidence that feedback locked waveforms were studied, producing 125 possible papers, of which four met the criteria for eligibility.

The reference lists of all eligible papers were checked, along with those of two recent reviews of the FRN (San Martin, 2012; Walsh & Anderson, 2012), producing one further eligible paper. In total, these search criteria resulted in the inclusion of 47 datasets from published papers in our meta-analysis.

Unpublished data. In an effort to include unpublished data, all first or corresponding authors of the selected papers were contacted with a request for unpublished data. A number of other researchers were also contacted, identified as follows. Papers returned by the searches described above which had been rejected were re-examined, and 154 authors added to a contact list. A search of theses using the ProQuest Dissertations and Theses database and the Ethos database returned 73 hits for the primary search string and 370 for the secondary one. The contents pages of these theses were read online and 17 authors added to the contact list. Abstracts of 56 conference papers, extracted from the searches described earlier, were read, and on this basis 8 more authors were added. In the course of contacting authors, a further four suggestions were garnered. 171 out of 183 email addresses were successfully obtained by internet search and these researchers contacted. Responses were obtained from 51

researchers. Four entirely unpublished datasets were retrieved by this process, and one dataset associated with a published paper in which the requisite waveforms had not been presented. Three unpublished studies of the author's own were also added. Therefore, we finally included 55 datasets into our meta-analysis, 47 from published data, 8 from unpublished data.

Validation data. As this is the first implementation of the GGA technique, we sought to validate it by comparing its findings with those resulting from conventional meta-analysis based on standardized effect sizes. For a meaningful comparison, it was important that these standardized effect sizes were generated in the same fixed interval of the waveform as that used for the GGA analysis. Effect sizes (or their derivatives) reported in the original papers did not correspond to this, or any fixed interval. It was their variability that prompted development of the GGA technique. To carry out the validation we therefore contacted authors of all the 55 papers used in our GGA analysis with a request for their original data so that we might calculate standardized effect sizes in the designated interval ourselves. This request returned original data for 14 of the 29 magnitude studies and 13 of the 26 likelihood studies. These studies are hereafter referred to as the validation dataset.

5.3.4 Coding Procedures: Generating Great Grand Averages Waveforms

Digitizing of published waveforms was performed with PlotDigitizer (<http://sourceforge.net/projects/plotdigitizer/>). Electronic copies of experiments were accessed, and the figures containing the requisite waveforms were enlarged and then opened in the PlotDigitizer software. Digitizing began by using a mouse to calibrate the minimum and maximum values of the x and y axis to the distance they occupied on the screen, thus defining the co-ordinate space of the area of the figure. The co-ordinates described by the actual ERPs were extracted by using a mouse to manually lay points along the waveforms at approximately 5 ms intervals. These were then run through a

purpose-written program that linearly interpolated co-ordinates at 1 ms intervals between the existing manually assigned ones. For every waveform undergoing the process, this generated a series of voltage values at discrete 1 ms intervals that made the subsequent process of averaging across studies tractable. The coordinates were immediately re-plotted to visually check that they corresponded to the original waveform they were taken from to prevent gross errors. All waveforms were digitized twice in this fashion, partly to improve accuracy and partly to allow reliability checks discussed below.

The consequence of this digitizing process was that for each study in the meta-analysis, we were able to recover the data that underlay the four relevant grand average waveforms, plus some measurement error. For the 27 experiments that were also represented in the validation dataset, original data replaced the digitized versions for the bulk of subsequent analysis. In these 27 cases the digitized versions were merely used to assess the degree of digitizing measurement error, as described later.

5.3.5 Coding Procedures: Quantifying the FRN

Figure 27 demonstrates the difference waveforms that lie at the heart of this meta-analysis, both analytically and conceptually, and this figure is referred to extensively throughout the chapter. The figure addresses an important issue of nomenclature, which was raised in Chapter 4. The FRN is typically operationalized as the difference between good (+RPE) and bad (–RPE) outcomes, i.e. a valence main effect. However, Holroyd and Coles claimed the FRN encoded a quantitative RPE, incorporating RPE size as well as valence, entailing a valence x RPE size interaction. To keep the distinction clear, we follow precedent in the present paper by using the term FRN to refer to a component responsive to the main effect of valence. We introduce the term RPE-FRN to refer to a component responsive to the interaction of valence and RPE size, the hypothesis under test in this meta-analysis. The simplest demonstration in

support of Holroyd and Coles would be a single component showing both such effects in the same interval. However it is also possible that the effects could be asynchronous, suggesting a quantitative RPE encoder of the kind envisaged by Holroyd and Coles accompanied by other components merely coding the sign of an RPE but not its size. In the simulated data of Figure 27, for example, the RPE-FRN in pane e occupies a briefer interval than the FRNs in panes c and d. This important distinction between FRN and RPE-FRN notwithstanding, at many points in the forthcoming discussion a point refers equally to both terms. Except in cases where we wish to make a point specific to the valence x RPE size interaction we refer simply to “the FRN”.

The grand average waveforms that were recovered by the digitizing process were submitted to the differencing process shown in Figure 27 in order to establish whether an RPE-FRN was present. As noted earlier, the RPE-FRN refers to a component responding to the interaction of RPE size and valence. Such an interaction is present when the difference waves shown in panes c and d of Figure 27 differ in amplitude. The effect size of the RPE-FRN component was thus the amplitude of the waveform corresponding to the difference of difference waves shown in pane e of Figure 27, and the significance of this effect size was based on a comparison of the amplitudes of its constituent difference waves, i.e. those corresponding to panes c and d, with this comparison made across the sample of either likelihood ($k = 26$) or magnitude ($k = 29$) studies.

Difference wave amplitude is typically measured either by using the waveform’s peak within a set interval, or its mean amplitude within a, usually smaller, interval. To provide a robust test of whether an RPE-FRN was present, we used both measures. The interval in which the measures were taken was determined by the average of those intervals used in the original papers. Those studies that used a mean amplitude measure

produced an average measurement interval of 228–344 ms, while those using a peak amplitude measure produced an average measurement interval of 128–460 ms.

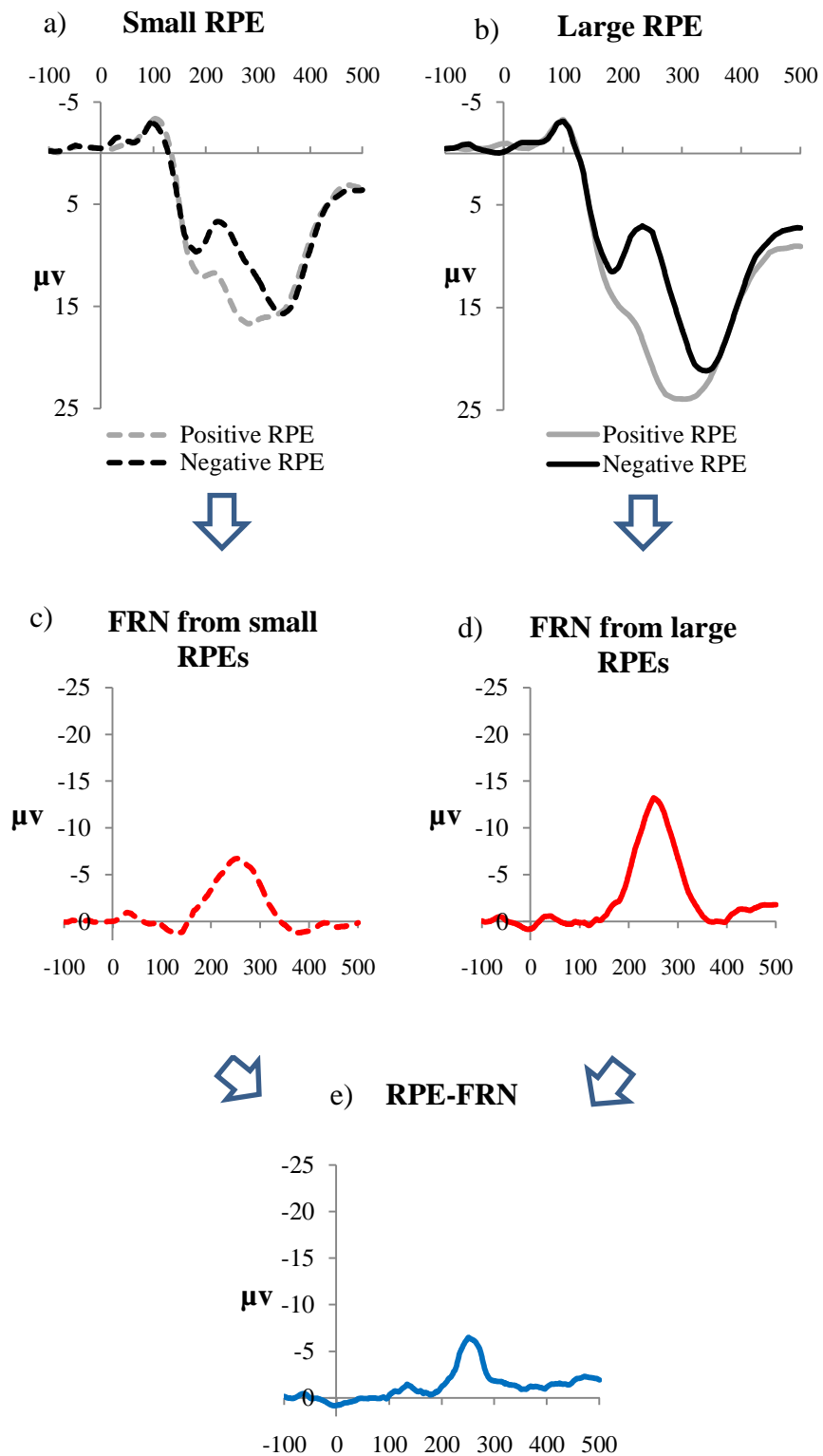


Figure 27 How the FRN is studied

How the FRN is studied. Panes a and b show grand average waveforms for four experimental conditions taken from a 2 x 2 design manipulating valence (+RPEs vs. -RPEs) and RPE size (large vs. small). A given experiment would manipulate RPE size using either outcome magnitude or likelihood. The simple waveforms in panes a and b show complex peaks which are the result of consecutive, overlapping components, many of which are unknown. Difference waves (FRNs), constructed by subtracting the +RPE outcome waveform from the -RPE outcome waveform are shown in panes c and d. These control for components unrelated to valence, allowing the valence effect to be more clearly seen. Comparison of panes c and d also suggests that the amplitude of the FRN in these data is sensitive to RPE size. This is definitively shown by differencing the difference waves in pane e. Collectively, the figures represent the prediction of Holroyd and Coles' theory, with pane a corresponding to low magnitude or likely outcomes and pane b to high magnitude, or unlikely outcomes

In addition to exploring the effect of magnitude and likelihood modulators on the FRN, we were interested in the effects of these variables in their own right, i.e. their main effects. To study these, rather than differencing the valence variable, it was collapsed out at each level of magnitude and likelihood, allowing the comparison of high and low magnitude waveforms and high and low likelihood waveforms. Thus in the scheme shown in Figure 27, an average waveform was created in each of panes a and b and these were then differenced (small RPE – large RPE) to produce an RPE size main effect difference wave.

5.3.6 Statistical Methods

Simple and standardized effect sizes. The differencing process described above was performed on each individual study, generating an effect size for the RPE-FRN, thus allowing a test for the significance of this effect size across the studies that made up the dataset. This process made no use of the standard deviation of the effect size *within a given study* however, i.e. calculated across the participants of that study, nor could it do so, since the digitizing process only had access to grand average waveforms. As noted, this does not prevent us testing for the significance of the effect across studies, but does prevent the relative weighting of individual studies based upon the variance of their data. This is generally used to down-weight the contribution from studies showing high variability on the basis that their estimate of the effect under question can be assumed to be less reliable. Conventional meta-analysis achieves this weighting up front by using *standardized effect sizes* (often referred to simply as “effect sizes”) as the unit of analysis, which down-weight effects when they are underlain by high variability. The standardized effect size metric in which this is most obviously expressed is Cohen’s *d*, which is the difference between two scores of interest divided by their pooled standard deviation. Standardized effect sizes can be contrasted with *simple effect sizes* (Baguley, 2009) or “raw mean differences” (Bond, Wiitala, &

Richard, 2003) which, as the name suggests, are equivalent to Cohen's d without any division by standard deviation. Simple effect sizes are what are produced by the GGA technique and what are used in the GGA meta-analysis presented here.

Both Baguley and Bond et al. have argued the virtues of working with simple effect sizes over standardized ones, noting the ease with which they can then be used to practically guide future studies (in the present case, a simple effect size informs researchers of the size in microvolts that they can expect to be working with, for example) and observing that the standard deviations that are used to calculate Cohen's d are themselves subject to the sampling error they purport to correct for. Another reason why standardized effect sizes have become the norm in meta-analyses is that they allow the comparison of scores derived from different scales of measurement, which is not an issue here, where the metric is always voltage. Nevertheless, the use of simple rather than standardized effect sizes is a notable feature of the GGA technique and we later examine its consequences using the validation dataset.

Testing the hypothesis I: T-tests on GGAs. Our hypothesis was that the FRN would be greater when RPEs were large rather than small, as described in panes c and d of Figure 27. Since the criterion for a generalized RPE encoder is that it should be modulated by both reward magnitude and likelihood, these two modulators were tested separately. In each case, a paired samples t-test was conducted of the amplitude of FRNs constructed from small RPEs vs. large RPEs. Four tests were done in total, on peak measures and mean amplitude measures of the magnitude and likelihood modulated FRNs. T-tests were entirely analogous to those which might be performed on individual FRN experiments but at "one level higher" using grand average data as data points, rather than participant average data.

Because sample size differed over studies, and conventional meta-analysis typically incorporates this information (Field & Gillett, 2010; Hunter & Schmidt, 2004),

weighted t-tests were used. The t statistic was calculated with the standard formula for paired samples

$$t = \frac{\bar{x}_D}{\frac{s_D}{\sqrt{N}}}$$

Where \bar{x}_D is the mean difference of the paired samples, and s_D its standard deviation.

However, \bar{x}_D was a *weighted* mean difference, calculated from k individual study mean differences (x) whose sample size was used as a weight (w), as follows

$$\bar{x}_D = \frac{\sum_{i=1}^k w_i x_i}{\sum_{i=1}^k w_i}$$

The standard deviation of this difference was also weighted, as follows

$$s_D = \frac{\sum_{i=1}^k w_i (x_i - \bar{x}_D)}{\sum_{i=1}^k w_i}$$

Unless otherwise stated, all statistics performed on GGAs used weighted means and standard deviations.

Sensitivity to publication bias was assessed by inspection of funnel plots followed by trim and fill (Duval & Tweedie, 2000) implemented in R using the metafor package (Viechtbauer, 2010).

Testing the hypothesis II: Data driven cluster randomization of GGAs. The analysis described above provides a fair but straightforward test of the hypothesis since the FRN was quantified in an interval determined *a priori* by the existing literature. However, it remains possible that this interval is a poor choice, certainly for capturing the RPE-FRN, i.e. the response to the interaction of RPE size and valence. We therefore also performed a secondary analysis using the cluster randomization procedure of Maris and Oostenveld (2007) described in Chapter 3. All procedures were equivalent to those previously used. This examined the full length of waveforms for evidence of an RPE-FRN component. As well as addressing the danger of using the wrong interval, this had

the secondary advantage that it could extract the observed interval of the RPE-FRN post hoc.

Heterogeneity of GGAs. Meta-analyses typically report heterogeneity, a measure of the likelihood that the sample effect sizes in the meta-analysis are drawn from more than one population. This is shown by a variance across sample effect sizes which exceeds that expected from the within-study variances. Since within-study variances are unknown under the GGA technique, heterogeneity cannot be measured. It can however be implied by demonstration of the significant effect of moderators.

Moderator analysis. This is conventionally performed in conjunction with a standardized effect size based meta-analysis, something we could not do with the GGAs, as we could not compute standardized effect sizes. To test for the effects of moderators, we performed univariate analyses with the moderator as a single categorical independent variable. The dependent variable was the simple effect size of the RPE-FRN. Unweighted effect sizes were used in an ANCOVA analysis with weighting applied using the weighted least squares function. To maximize the power of the moderator analysis, likelihood and magnitude modulated studies were analyzed together. Since validation of the GGA technique (reported later) suggested that mean amplitude measures produced closer estimates to an ideal conventional meta-analysis than peak measures, only mean amplitude measures were used for moderator analysis. Confounding of moderators was checked using contingency coefficients of all possible pairs of the four moderators, and where significant χ^2 values were found, entering the confounding moderators as covariates.

Meta-analysis of validation data. Conventional meta-analysis was performed using standardized effect sizes of the RPE-FRN generated from original data obtained from authors. Differencing of waveforms and calculation of t values was performed in the same way as was done for GGAs, with t values then converted to Cohen's d . A

calculation from t values was used rather than direct calculation using the mean difference divided by its standard deviation because of problems arising from the standard deviation term of paired samples designs. As Dunlap, Cortina, Vaslow, and Burke (1996) have observed, paired samples designs increase power by reducing the standard deviation term. This makes it easier to detect an effect (t is increased, for example). However, the paired samples design does not change the effect's size, which is what d purports to represent. Using the paired samples standard deviation in calculating d therefore conflates effect size with effect significance and inflates the estimate of d . Since the degree of this inflated estimate is proportional to the additional power the paired design provides, and this in turn is proportional to the extent to which the paired scores move together, d can be corrected by using the correlation coefficient of the two conditions underlying the t -test. Dunlap et al.'s formula for this unbiased calculation is shown below and was used for calculation of d . Note that the r term should not be confused with an effect size metric.

$$d = t. \sqrt{\frac{2(1 - r)}{n}}$$

Meta-analysis was conducted using the method of Hunter and Schmidt (2004), with studies weighted by their sample size rather than inverse variance, since this allowed the closest comparison with the GGA technique. A random effects model was used, due to concerns over the generalizability of fixed effects models (Field & Gillett, 2010). The meta-analysis produced an estimated effect size, confidence intervals for this estimate, and, most importantly for our validation purposes, a significance test that could be compared to that produced by the GGA technique. Heterogeneity was measured using the Q statistic. Analyses were implemented in the macros provided by Field and Gillett (2010) apart from trim and fill which was implemented in R using the metafor package.

Meta-analysis of published data. While the GGA technique is premised on the unsuitability of published FRN effect sizes for meta-analysis, we ran a further meta-analysis using published effects for illustrative purposes. The effect size measure used was once again Cohen's *d*. Effect sizes were frequently not reported in the published papers, and where they were it was typically in the form of partial eta squared. Values of *d* were therefore calculated directly from reported test statistics using conventional approximations. Where the reported statistic was *t*, the Dunlap formula above was used with *r* estimated at 0.5: the average correlation found in our validation dataset was in fact 0.49. Where the statistic given was an *F* value, that is, rather than a difference of difference waves, the RPE-FRN effect size was expressed as a valence x RPE size interaction, Rosenthal's (1991) conversion was used:

$$d = 2 \left(\sqrt{\frac{F}{df_d}} \right)$$

In cases where effects were reported as “non-significant” or an inequality based on a canonical value such as $F < 1$ was given, *d* was set to zero. If a non-canonical value of a statistic or *p* value was given (e.g. $p < .06$) this was taken as the actual value. Meta-analysis was then performed as described for the validation data.

5.4 Results

5.4.1 Modulation of the FRN by Magnitude and Likelihood.

Figure 28 shows simple great grand average waveforms for magnitude and likelihood designs. Figure 29 depicts the central test of the hypothesis. It can be seen from Figure 29a that the FRN for high magnitude outcomes is of greater amplitude than the FRN for low magnitude outcomes, suggesting that the FRN is sensitive to outcome magnitude in the manner predicted. This sensitivity is plotted as an RPE-FRN, that is, the difference of the high magnitude difference wave and the low magnitude difference

wave. A paired samples t-test on mean FRN amplitudes in the interval 228–334 ms revealed a significant difference ($M_{\text{low}} = -1.52 \mu\text{v}$, $M_{\text{high}} = -2.20 \mu\text{v}$, RPE-FRN simple effect size = $-.68 \mu\text{v}$, $t(28) = -4.41$, $p < .001$). A t-test on peak FRN amplitudes in the interval 129–447 ms also showed a significant difference ($M_{\text{low}} = -2.30 \mu\text{v}$, $M_{\text{high}} = -3.11 \mu\text{v}$, RPE-FRN simple effect size = $-0.81 \mu\text{v}$, $t(28) = -3.11$).

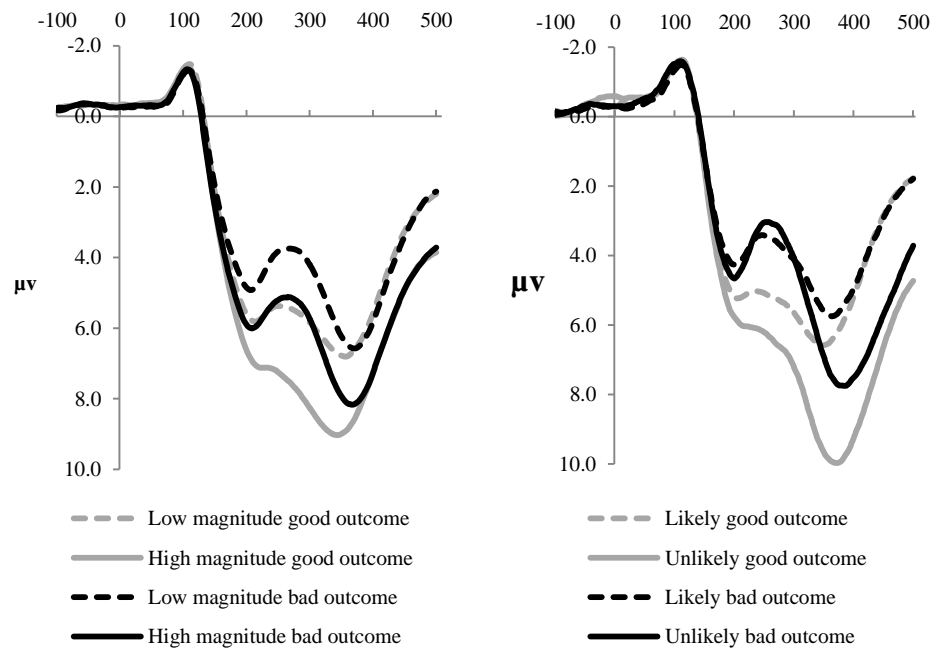


Figure 28 Simple GGA waveforms

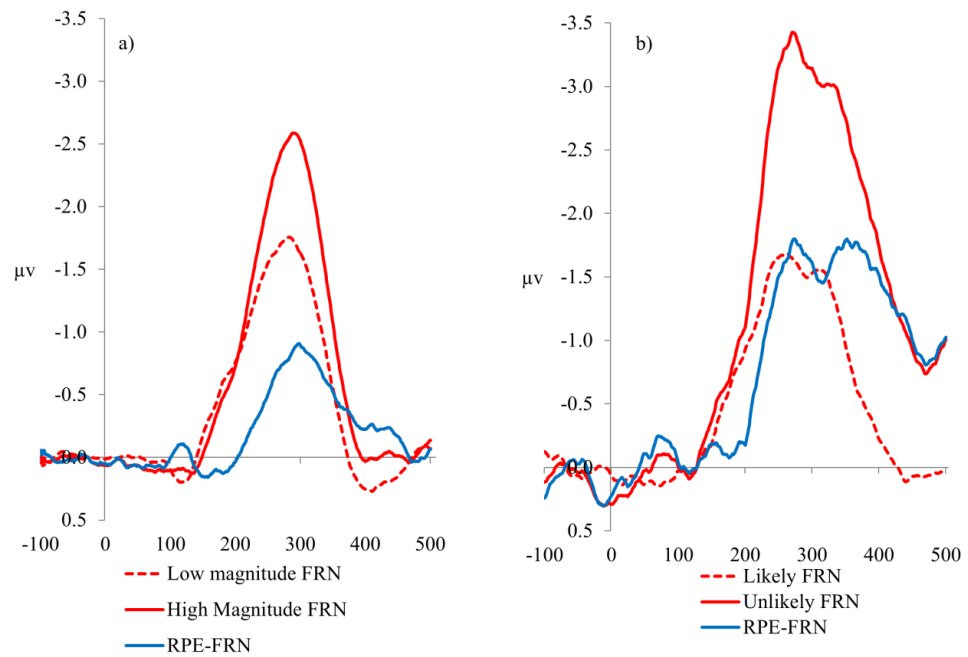


Figure 29 FRN GGAs showing effect of RPE-FRN

Modulation of the FRN by a) magnitude, and b) likelihood. Difference waves (FRNs) are created from $-RPE$ minus $+RPE$ waveforms. The RPE-FRN simple effect size is the difference of the two difference waves

Similar comparisons for the likelihood modulator can be seen in Figure 29b, where it can be seen that, as predicted, the FRN for unlikely outcomes is of greater amplitude than the FRN for likely outcomes, again generating an RPE-FRN. The effect was significant under a mean amplitude measure in the interval 228–334 ms ($M_{\text{likely}} = -1.56 \mu\text{v}$, $M_{\text{unlikely}} = -3.10 \mu\text{v}$, RPE-FRN simple effect size = $-1.54 \mu\text{v}$, $t(25) = -5.44$, $p < .001$) and a peak measure in the interval 129–447 ms ($M_{\text{likely}} = -2.84 \mu\text{v}$, $M_{\text{unlikely}} = -4.65 \mu\text{v}$, RPE-FRN simple effect size = $1.84 \mu\text{v}$, $t(25) = -5.62$, $p < .001$). The RPE-FRN simple effect sizes for both modulators under the mean amplitude measure are shown as a forest plot in Figure 30. As a further check, the t-tests described above were conducted on unweighted scores to ensure that the effects were not unduly affected by a few studies with large sample sizes. All effects remained strongly significant.

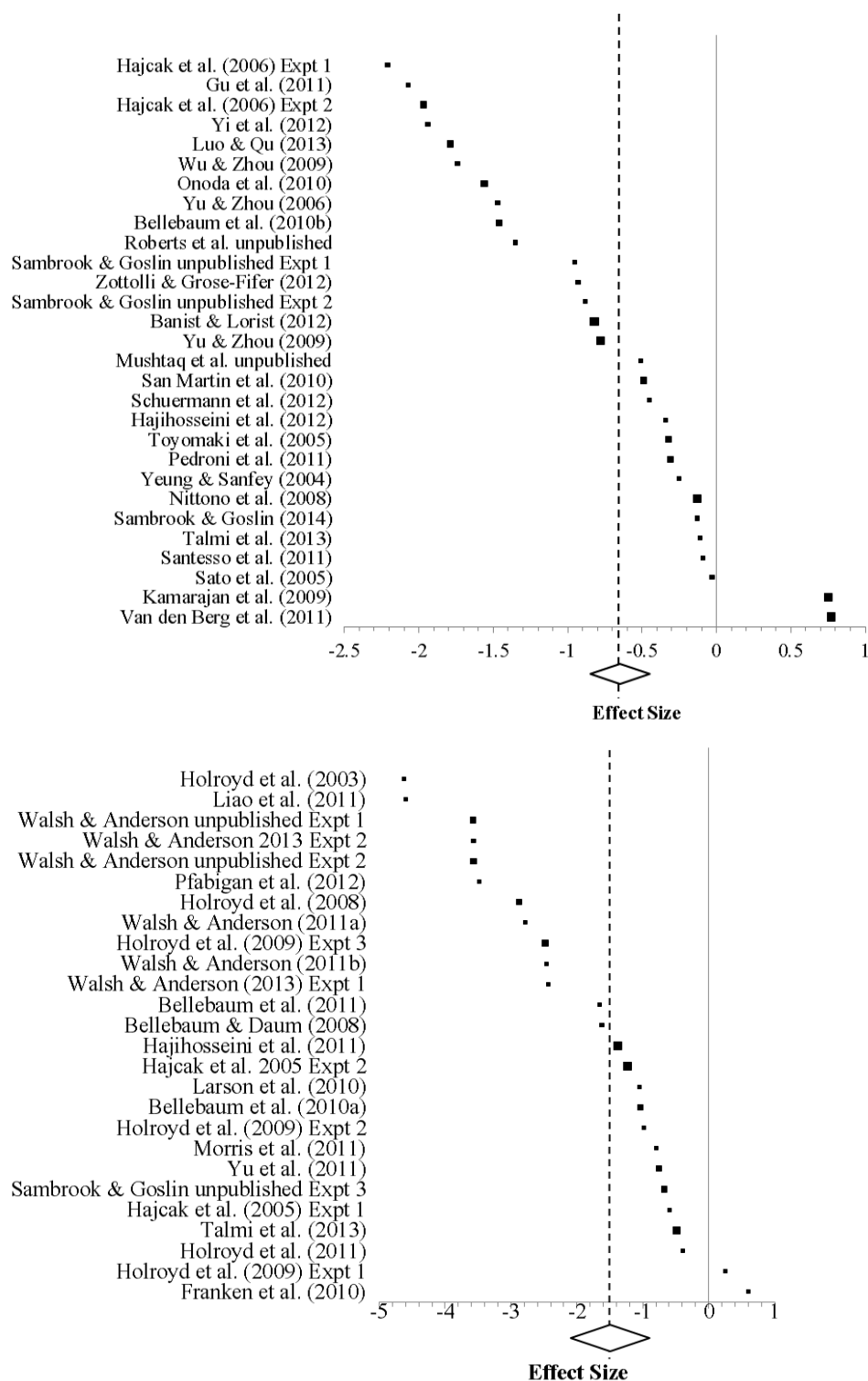


Figure 30 Forest plot showing RPE-FRN simple effect size

Forest plot showing RPE-FRN simple effect size in a) magnitude and, b) likelihood designs. The size of squares indicates the sample size, which also constituted the

weighting in the GGA meta-analysis. The diamond shows average weighted simple effect size and 95% confidence intervals.

The hypothesis was thus supported using a quantification of the FRN based on *a priori* intervals derived from the literature. The Maris and Oostenveld procedure was then used to more accurately determine the latency of the RPE-FRN specifically. For the magnitude modulator, a single significant cluster of RPE-FRN activity was found (Monte Carlo $p = .0001$), running from 240–341 ms, with the effect greatest at 298 ms ($-.91 \mu\text{v}$). For the likelihood modulator a single cluster of RPE-FRN activity was found (Monte Carlo $p < .0001$), running from 209 ms to the edge of the measurement interval at 500 ms. The effect was equally great at 274 and 352 ms ($-1.80 \mu\text{v}$) but much more significantly so at the earlier peak: $t(25) = 6.46$.

5.4.2 Publication Bias

Publication bias was assessed by inspection of the funnel plots shown in Figure 31. Because these suggested a small degree of asymmetry, albeit largely among studies with large rather than small sample sizes, we applied a trim and fill procedure. This was implemented by entering the simple effect sizes derived from GGA analyses into a conventional meta-analysis, rebalancing potential asymmetry in the funnel plots by adding additional imputed studies, and then recalculating effect sizes. In fact, this procedure resulted in no additional studies being imputed, leaving effect sizes unchanged, and demonstrating absence of publication bias

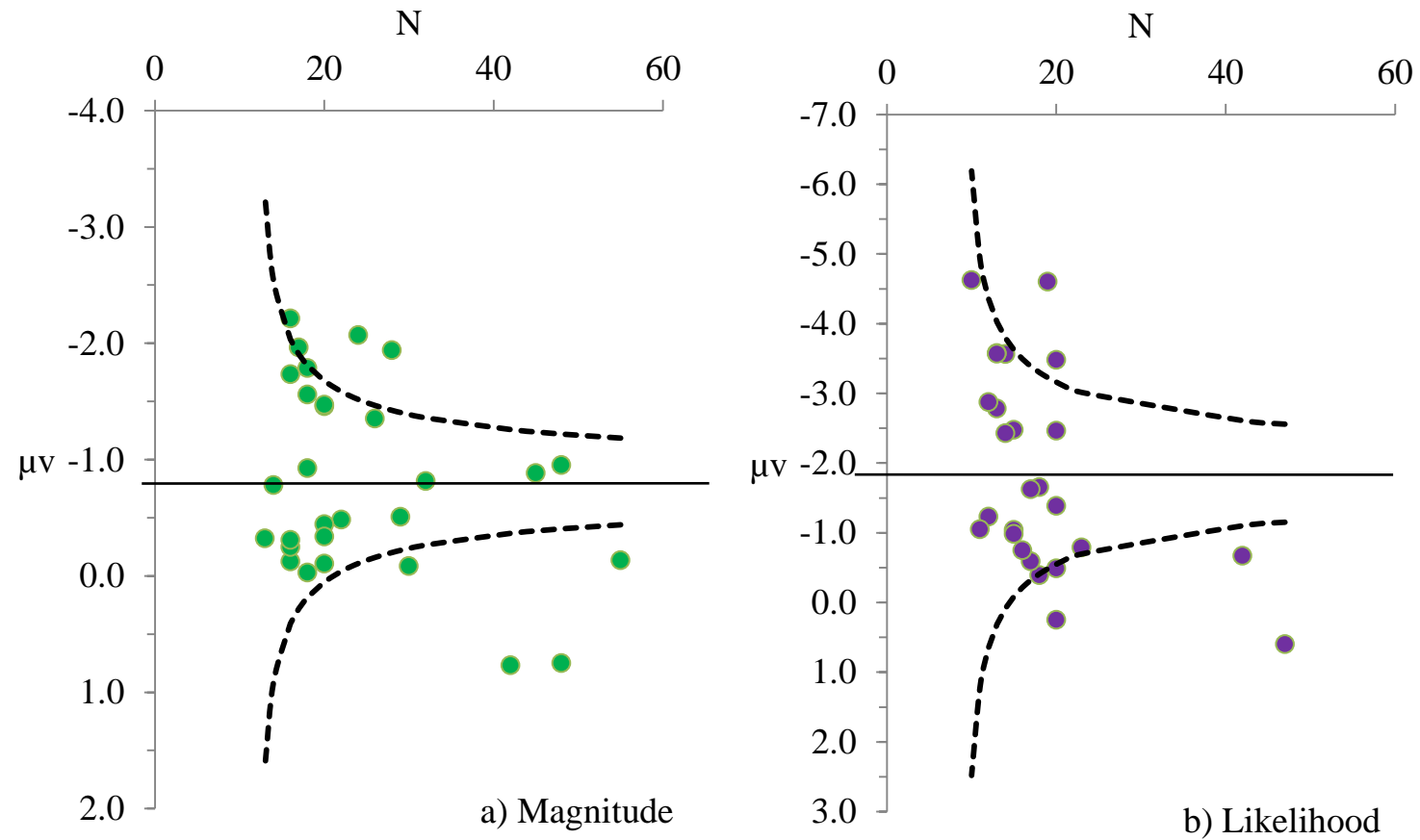


Figure 31 Funnel plots for the unweighted simple effect size of the RPE-FRN

Funnel plots for the unweighted simple effect size of the RPE-FRN under a) magnitude, b) likelihood. Dotted lines represent three standard deviations.

5.4.3 Moderator Analyses of the FRN

χ^2 tests revealed strong associations ($p < .001$) between three of the four moderators: modulator, domain and control over outcome. To test the effect of each moderator individually, while controlling for the effects of the others, analysis of covariance was used with the confounding moderators entered as covariates. Once again, a mean amplitude measure in the interval 228–334 ms was used.

Modulator. No significant effect of modulator on RPE-FRN simple effect size was found (magnitude: $-.72 \mu\text{v}$, $k = 27$; likelihood: $-1.60 \mu\text{v}$, $k = 24$; $F_{1,47} = 2.92$, $p = .09$). The apparent strong effect of modulator shown by a comparison of the subplots in Figure 29, and the means above, was due to the mediating effect of control over outcome (see below). Because Figure 29 also suggested the possibility that the RPE-FRN of likelihood experiments occupied a longer interval than that for magnitude experiments, this was investigated using a mean amplitude measure in the interval 335–500 ms. The effect of modulator on RPE-FRN in this interval proved to be narrowly non-significant (magnitude: $-.23 \mu\text{v}$, $k = 27$; likelihood: $-1.38 \mu\text{v}$, $k = 24$; $F_{1,47} = 3.80$, $p = .057$).

Control over outcome. A significant effect of control over outcome was found, with RPE-FRN amplitude increasing as control grew (passive: $-.07 \mu\text{v}$, $k = 5$; guessing: $-.88 \mu\text{v}$, $k = 34$; rule implementation: $-2.47 \mu\text{v}$, $k = 12$; $F_{2,46} = 9.71$, $p < .001$). Post hoc comparisons revealed all pairwise comparisons to be significant ($p < .05$). A significant effect was also found in the later interval of 335–500 ms (passive: $-.41 \mu\text{v}$, $k = 5$; guessing: $-.25 \mu\text{v}$, $k = 34$; rule implementation: $-2.56 \mu\text{v}$, $k = 12$; $F_{2,46} = 7.40$, $p = .002$). Post hoc comparisons in this interval revealed that rule implementation produced a significantly stronger RPE-FRN than passive or guess designs ($p < .05$), but these two

levels did not significantly differ. Waveforms of the RPE-FRN for the three levels (with modulator collapsed out) are shown in Figure 32.

Domain. No effect of domain was found (gain: $-1.37 \mu\text{v}$, $k = 25$; mixed: $-.82 \mu\text{v}$, $k = 26$; $F_{1,51} < .01$)

Magnitude Cueing. No effect of magnitude cuing on the RPE-FRN was found (cued: $-.70 \mu\text{v}$, $k = 20$; un-cued: $-.59 \mu\text{v}$, $k = 8$, $F_{1,26} < .1$)

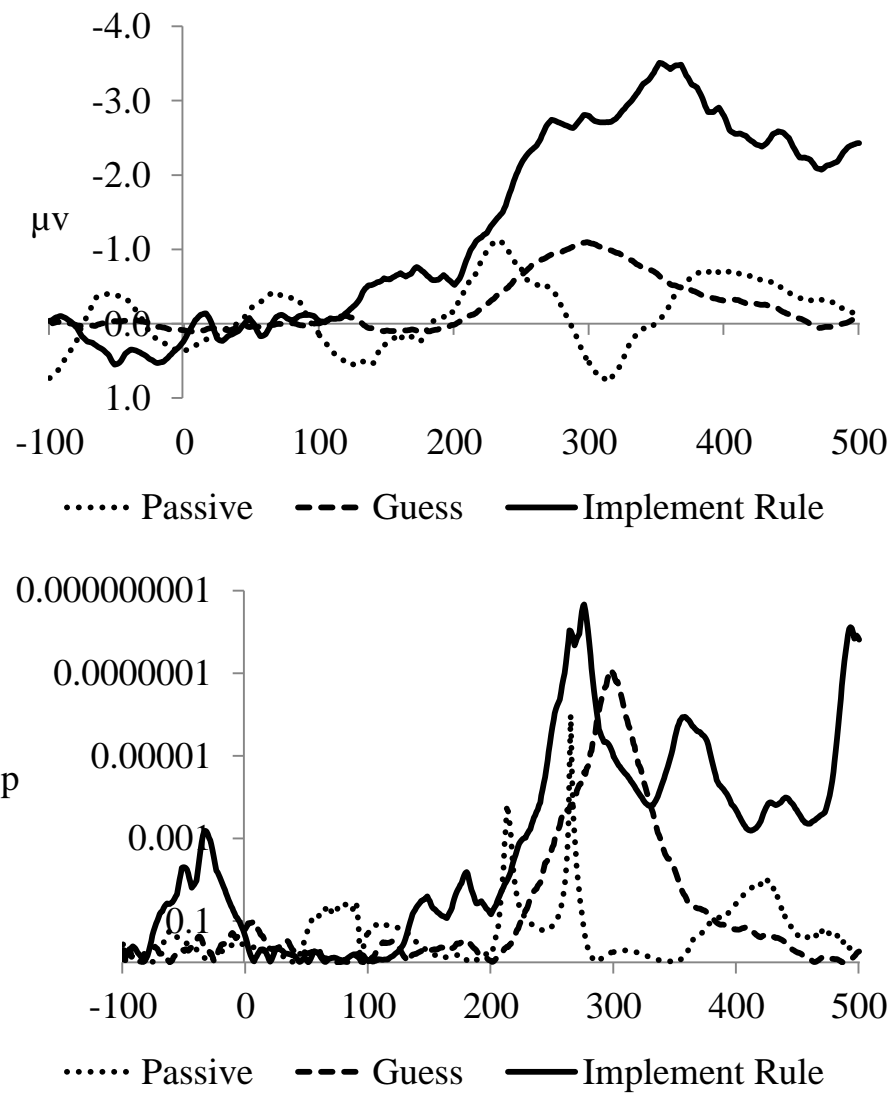


Figure 32 Moderator analysis of the FRN

RPE-FRN at different levels of the “control over outcome” moderator: a) simple effect size, b) significance of simple effect size

5.4.4 Validation of the GGA Technique

Where an electrophysiological component is quantified in diverse ways in a literature, we have argued that the GGA technique is superior to conventional meta-analysis because it allows quantification to be made in a standardized interval. Nevertheless, the GGA technique suffers two potential drawbacks relative to conventional meta-analysis. The first is that the process of recovering original data from published figures introduces measurement error. The second is that the GGA technique has no access to information on within-study variability and treats each study as equivalent in this regard. In comparison, conventional meta-analysis uses standardized effect sizes which incorporate a measure of this variability, serving to down-weight effects found in studies with high variability. The output of the GGA analyses was therefore compared to the output from analyses performed on the original data obtained directly from authors, allowing us to assess the impact of these potential drawbacks.

Digitizing error. Digitizing error could be easily measured by comparing the digitized data with the original data in the 27 studies of the validation dataset. The difference between the two data sources could either be as a result of the process used to digitize the figures, or discrepancies between the original data and the figures used in publication. To quantify the digitizing error, a second coder (naive to the hypothesis under test) repeated the digitizing process of the original coder for the whole of the validation dataset. This allowed us to assess the degree of error within a single coder (intra-coder error), between the two coders (inter-coder error), and between the main coder and the original data (coder-original error). Average errors for the RPE-FRN in the critical interval 228–334 ms were as follows. The main coder showed an intra-coder error of $-.011 \mu\text{v}$ ($SD = .099$), and the secondary coder $-.004 \mu\text{v}$ ($SD = .028$).

Comparison of the two coders' average scores revealed an inter-coder error of $-.005 \mu\text{v}$

($SD = .075$). Comparison of the main coder with original data revealed a coder-original error of $.096 \mu v$ ($SD = .327$). Intra- and inter-coder error was very low suggesting that an accurate digitizing of a published figure is unproblematic. Error rates between the main coder and the original data were higher than between the two coders, implying that the main source of discrepancy lies in the manner with which the original authors have prepared graphs for publication. Examination on a case by case basis revealed that this was isolated to a few studies that would appear to have used a low-pass filter on the figure, but not the data, or a degree of erroneous vertical or horizontal translation of the waveform of one of the experiment's conditions. However, the amount of coder-original error is nevertheless very modest compared to the average simple effect sizes found in the GGA meta-analysis. Furthermore, it should be stressed that these digitising errors did not affect the statistical testing applied to GGAs earlier, since original data was used in their stead. They merely give an estimate of the extent of the error in the remaining 27 studies for which no original data was available, and for the use of the technique generally.

Meta-analysis of original data. To assess the overall performance of the GGA technique, an “ideal” conventional meta-analysis was conducted using the same interval as used for the GGA meta-analysis, but with standardized effect sizes calculated from the original data in the validation dataset. The results of this meta-analysis were then compared to a GGA meta-analysis run on the same subsample of 27 studies. Results of both meta-analyses are given in Table 3. Quite aside from its role in validating the GGA technique it can be seen that the conventional meta-analysis also strongly supports this study's hypotheses, showing a significant RPE-FRN effect size under both the magnitude and likelihood modulator. With regard to validating the GGA technique here and more generally, it can be seen that the two meta-analytic methods give comparable results in regard to significance testing of the mean amplitude measure. The z statistic

from conventional meta-analysis and the t statistic from the GGA technique are very similar under both magnitude (5.36 vs 5.27) and likelihood (6.55 vs 6.19) modulators. For the peak measure, the ideal conventional meta-analysis reveals the GGA technique to have been conservative. This is to be expected, as the GGA technique measures the peak amplitude of grand averages rather than participant averages, and thus is subject to greater temporal smearing due to individual differences in latency across participants. Note that while Table 4 reports both average standardized effect size under ideal conventional meta-analysis and average simple effect size under the GGA technique, these should not be directly compared as they are denominated in different units. For GGAs they are measured in microvolts, for the conventional meta-analysis, in standard deviations of microvolts. Effect sizes for individual studies are given in Table 5. Note also that the validation dataset can be considered representative insofar as there was no significant difference in the RPE-FRN simple effect size of studies in or out of the validation dataset, $t(53) = 1.54$, $p = .13$.

Table 4 Results of a conventional vs. GGA meta-analysis

Conventional Meta-analysis on validation set original data						GGA Meta-Analysis on validation set	
Modulator	Measure	Average standardised effect size	Significance	Heterogeneity	Effect after trim and fill	Average simple effect size (μv)	Significance
<i>Magnitude</i>	Mean	-.49 [-.67,-.31]	$z = 5.36, p < .001$	$\chi^2 (13) = 11.75, p > .05$	-.49	-.79 [-1.06, -.42]	$t = 5.27, p < .001$
	Peak	-.51 [-.68,-.34]	$z = 6.01, p < .001$	$\chi^2 (13) = 12.43, p > .05$	-.51	-.89 [-1.34, -.43]	$t = 4.12, p < .001$
<i>Likelihood</i>	Mean	-.71 [-.92,-.50]	$z = 6.55, p < .001$	$\chi^2 (12) = 12.00, p > .05$	-.71	-2.17 [-2.75, -1.59]	$t = 6.19, p < .001$
	Peak	-.87 [-1.09,-.64]	$z = 7.43, p < .001$	$\chi^2 (12) = 12.90, p > .05$	-.87	-2.31 [-3.09, -1.52]	$t = 4.97, p < .001$

Table 5 Effect sizes in the validation data set

Effect sizes in the validation data set. Standardized effect sizes of the RPE-FRN are taken from original data and simple effect sizes of the RPE-FRN are taken from digitized waveforms

Study	Standardized effect size	Simple effect size (μv)
<i>Likelihood experiments</i>		
Bellebaum & Daum (2008)	-0.55	-1.49
Bellebaum et al. (2010a)	-0.26	-1.02
Bellebaum et al. (2011)	-0.46	-1.98
Hajihosseini et al. (2012)	-0.52	-1.57
Pfabigan et al. (2012)	-0.76	-3.28
Sambrook & Goslin unpublished Expt 3	-0.32	-0.69
Talmi et al. (2013)	-0.44	-0.53
Walsh & Anderson (2011a)	-0.78	-3.02
Walsh & Anderson (2011b)	-0.91	-3.20
Walsh & Anderson (2013) Expt 1	-1.01	-3.57
Walsh & Anderson (2013) Expt 2	-1.51	-4.43
Walsh & Anderson unpublished Expt 1	-1.42	-3.60
Walsh & Anderson unpublished Expt 2	-1.19	-3.63

Magnitude experiments

Banist & Lorist (2012)	-0.47	-0.84
Bellebaum et al. (2010b)	-0.34	-1.50
Gu et al. (2011)	-0.53	-1.84
Hajihosseini et al. (2012)	-0.14	-0.37
Luo & Qu (2013)	-0.44	-1.54
Onoda et al. (2010)	-0.36	-0.43
Pedroni et al. (2011)	-0.26	-0.30
Roberts et al. unpublished	-0.50	-1.31
Sambrook & Goslin unpublished Expt 1	-1.23	-0.78
Sambrook & Goslin unpublished Expt 2	-0.78	-0.89
Sambrook & Goslin (2014)	-0.27	-0.13
Santesso et al. (2011)	-0.25	-0.31
Talmi et al. (2013)	-0.10	-0.08
Wu & Zhou (2009)	-0.78	-1.77

Meta-analysis of published effect sizes. We also performed a conventional meta-analysis of published effects. As previously stated, we believe this is an unsound meta-analysis because it draws on effect sizes measured in different intervals and from quite different quantifications of the FRN (e.g. mean amplitude, peak of difference wave, peak to peak of simple waves). Nevertheless it is interesting for comparative purposes and furthermore permits a quantifying of the simpler “face value” of accumulated reporting findings regarding likelihood and magnitude modulators. The meta-analysis was performed on a reduced dataset because a number of papers did not report statistics for the RPE-FRN effect (see Table 3) The average standardized effect

size for the magnitude modulator ($k = 15$) was non-significantly different from zero ($d = -.26 [-.80, .29]$, $z = .914$, $p = .361$). The average standardized effect size for the likelihood modulator ($k = 18$) was however significant ($d = -.95 [-1.34, -.56]$, $z = 4.82$, $p < .001$). Standardized effect sizes for individual studies are given in Table 6.

Table 6 Effect sizes of the RPE-FRN calculated from published data

Experiment	Statistic	N	Value	Cohen's d
<i>Likelihood experiments</i>				
Bellebaum & Daum (2008)	F	17	-6.41	-1.27
Bellebaum et al. (2010a)	F	15	0.00	0.00
Bellebaum et al. (2011)	t	18	-2.68	-0.63
Hajcak et al. (2005) Expt 1	F	17	0.00	0.00
Hajcak et al. (2005) Expt 2	F	12	0.00	0.00
Hajihosseini et al. (2012)	F	20	-5.00	-1.03
Holroyd et al. (2003)	t	10	-3.10	-0.98
Holroyd et al. (2009) Expt 1	t	20	-3.30	-0.74
Holroyd et al. (2009) Expt 2	t	15	-2.30	-0.59
Holroyd et al. (2009) Expt 3	t	15	-6.10	-1.58
Holroyd et al. (2011)	t	18	-2.30	-0.54
Liao et al. (2011)	F	19	-21.16	-2.46
Morris et al. (2011)	F	23	0.00	0.00
Pfabigan et al. (2012)	F	20	0.00	0.00
Walsh & Anderson (2011b)	F	20	-28.50	-2.45
Walsh & Anderson (2011a)	t	13	-3.62	-1.00
Walsh & Anderson (2013) Expt 1	F	14	-19.44	-2.45

Walsh & Anderson (2013) Expt 2	F	14	-14.26	-2.09
--------------------------------	---	----	--------	-------

Magnitude experiments

Banis & Lorist (2012)	F	32	6.20	0.91
Gu et al. (2011)	F	24	-17.02	-1.72
Hajcak et al. (2006) Expt 1	F	16	-1.78	-0.69
Hajihosseini et al. (2012)	F	20	-0.10	-0.15
Kamarajan et al. (2009)	F	48	74.72	2.52
Luo & Qu (2013)	F	18	-5.30	-1.12
Onoda et al. (2010)	t	17	-2.71	-0.66
Santesso et al. (2011)	F	30	7.55	0.82
Sato et al. (2005)	F	18	0.00	0.00
Schuermann et al. (2012)	F	20	-10.12	-1.46
Wu & Zhou (2009)	F	16	0.00	0.00
Yi et al. (2012)	F	28	-0.03	-0.04
Yu & Zhou (2006)	F	20	0.00	0.00
Yu & Zhou (2009)	F	14	-1.69	-0.72
Zottoli & Grose-Fifer (2012)	F	18	-23.60	-1.67

5.4.5 Main Effects of Magnitude and Likelihood

While the principal objective of the study was to test for the existence of an RPE-FRN by examining the FRN's sensitivity to modulation by magnitude and likelihood, a consideration of these modulators' main effects is also valuable in interpreting the post-feedback waveform that FRN studies are likely to generate.

Component overlap is an ever-present concern in ERP experiments and we felt it was very possible that an RPE-FRN would be superimposed on other components responding to magnitude, likelihood, or indeed valence, alone. Figure 33 represents all main effects in the form of difference waves. The RPE-FRN, calculated from magnitude and likelihood studies combined, is added for the purposes of comparison. Significance of main effects was determined using the Maris and Oostenveld technique. This revealed a magnitude main effect (Monte Carlo $p < .0001$), such that low magnitude outcomes were associated with a relative negativity in an interval running from 124 ms to the measurement boundary of 500 ms, with the effect greatest at 322 ms ($-2.10 \mu\text{v}$). Also revealed was a significant main effect of likelihood (Monte Carlo $p < .0001$), such that high likelihood outcomes were associated with a relative negativity in an interval running from 299 ms to the measurement boundary of 500 ms (Monte Carlo $p < .0001$), with the effect strongest at 426 ms ($-3.51 \mu\text{v}$). Finally there was a main effect of valence (Monte Carlo $p < .0001$), i.e. an FRN, in the interval 150–401 ms, with the effect greatest at 276 ms ($-2.27 \mu\text{v}$).

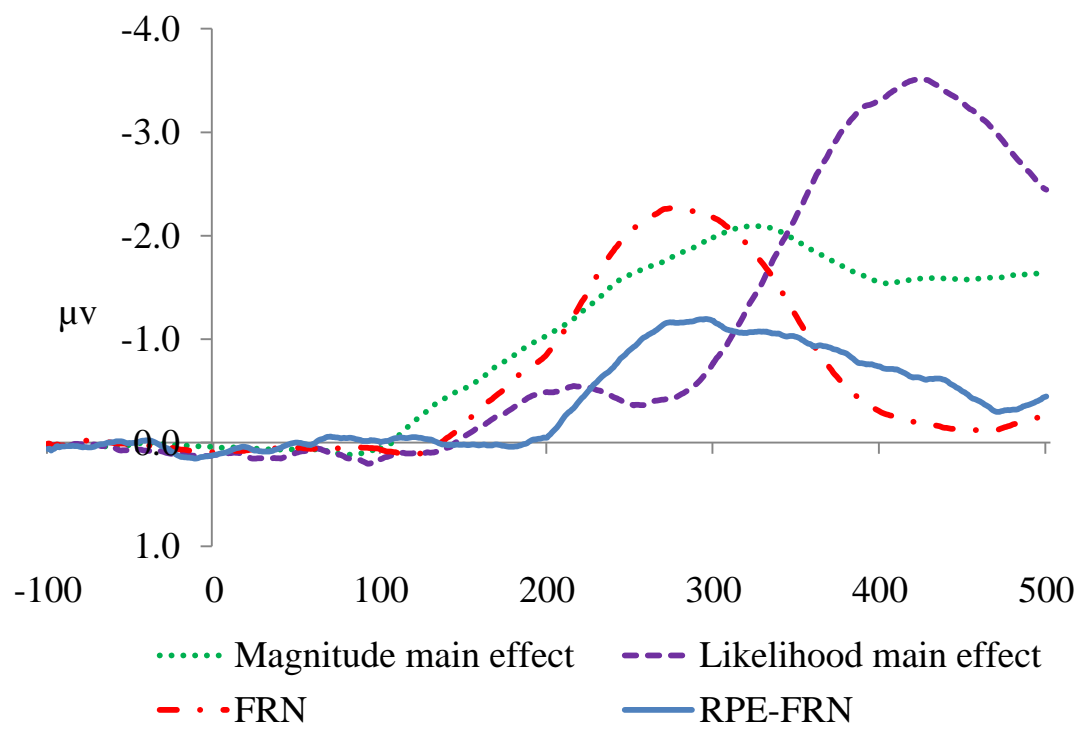


Figure 33 Main effects of magnitude, likelihood and valence (RPE-FRN also)

5.5 Discussion

5.5.1 The RPE-FRN and Main Effects of Valence, Magnitude and Likelihood

Holroyd and Coles (2002) proposed that the FRN encoded an RPE. The axiomatic model of RPEs of Caplin and Dean (2008) requires that the FRN show an interaction of RPE size and RPE valence if it is to constitute an RPE encoder. The present results are consistent with this requirement. FRNs created from large RPEs were of greater amplitude than those from small RPEs, both when RPE size was modulated by magnitude, and by likelihood. The demonstration that the FRN is responsive to variations in magnitude is important because it is a key requirement of a general RPE encoder, and evidence in previous experiments has largely been against this. The present meta-analysis shows that once quantification of the FRN is standardized, a clear magnitude effect on the FRN can be seen. It should be noted that the present meta-analysis does not address all of the axioms of the axiomatic model. In particular, the demonstrated interaction of RPE size and valence that lies at the heart of the meta-analysis might still have been produced by voltage shifts in response to increasing RPE size that are *of the same polarity* for both +RPEs and –RPEs, and which simply differ in their degree of response. That is to say, it may have been created, at least in part by an asymmetrical salience response of the kind shown by many factors in Chapter 4. The meta-analysis does not have access to the methods that were used to eliminate this possibility in Chapter 4.

A number of recent papers have reported evidence consistent with the FRN constituting a salience encoding (Hauser et al., 2014; Oliveira et al., 2007; Tobler et al., 2005) and it certainly appears that salience is coded in the post-feedback waveforms, as shown by the strong main effects of likelihood and magnitude in Figure 33, with these

main effects approximately twice the size of the RPE-FRN to which each modulator contributes. Strictly, such a salience encoder should show a strong main effect of RPE size (i.e. of likelihood and magnitude) but no main effect of valence, and no interaction of RPE size and valence (i.e. no RPE-FRN), since unsigned prediction errors should be insensitive to valence. However, once one allows for the possibility of asymmetrical salience encoders then the characterization of the FRN becomes more nuanced. We simply note that the present meta-analysis provided increased support for the claim that the FRN is an RPE encoder, in that an attempted falsification (notably in the case of the magnitude modulator) did not succeed.

In either case, it is clearly the case that multiple components contribute to activity at frontocentral electrodes, and this touches on an important conceptual point. This meta-analysis shows that frontocentral activity in the interval in which the FRN is typically measured is responsive to the main effects of magnitude, likelihood and valence, and also to the interaction of valence with both magnitude and likelihood. In keeping with the demonstrations in Chapters 3 and 4, it appears that multiple components operate in this interval. However the debate following the publication of Holroyd and Coles' theory has crystallized around the idea of a single component in this interval, whose character will ultimately be resolved through careful experimentation. In practice, we suspect that the character of the component described by a given experiment as the "FRN" will depend strongly on the interval in which this component is measured. For example, in Talmi et al.'s and Hauser et al.'s papers, evidence was presented favouring a salience account, however measurement was made at the latency of maximal FRN amplitude, that is the maximal main effect of valence (albeit that the interaction term was itself later analysed). While this was a pragmatic choice and based on precedent, this latency was nevertheless not necessarily one best suited to demonstrating an RPE-FRN if there was one to be found, since that is shown by a

valence x RPE size interaction, not a valence main effect. In practice, this resulted in these papers measuring effects at ~220 ms. This was prior to the period where the RPE-FRN was observed in this meta-analysis but where salience effects were marked in magnitude and close to significance for likelihood. In the FRN debate generally, we suspect that the sensitivity of the FRN to the key factors that are used to infer its function has depended on the latency of its measurement to a degree which has not been fully appreciated.

It is possible that in the future, separation of these components may be assisted by improved knowledge about their scalp distributions. Because of the limited and variable electrode arrays available, the present meta-analysis cannot offer guidance here. Furthermore, the FRN itself is partly defined by being maximal over frontocentral sites. Given that is a now well-established definition, it is likely any published example of the FRN would also have to demonstrate a frontocentral maximum, in order to be accepted as such. Therefore any meta-analysis of the FRN would be very likely reflect this established scalp distribution. In contrast, it seems likely that the later strong likelihood effect, and possibly magnitude effect shown in Figure 33 are P3 effects and would be maximal at more parietal locations.

The RPE-FRN was stronger when participants were engaged in a task over which they had reason to believe they enjoyed some control, a finding that reflects the increased effect size found in Chapter 4 relative to Chapter 3. In the strongest case of control, where participants implemented a known rule, the RPE-FRN also lasted much longer, as can be seen in Figure 32. These results suggest the possibility that the RPE-FRN might be selectively recruited by the apparatus of instrumental conditioning, rather than acting as a general purpose representation of value. Some caution must be exercised in regard to this finding, first, because subjective involvement was probably lower with reduced control (Yeung et al., 2005) and, second, because ten of the twelve

studies used for the “rule implementation” level of this moderator came from experiments conducted by just two authors.

The RPE-FRN in magnitude studies was unaffected by whether the magnitude of the forthcoming outcome was cued in advance. As such, it appears that the RPE-FRN does not scale RPEs to the range of outcomes on a given trial. We do not believe this should be regarded as evidence that RPEs are genuinely coded on an absolute scale however, since this would be functionally extremely limited and is biologically implausible. Scaling, or “adaptation”, is a ubiquitous feature of sensory processes, allowing, for example, the eye to discriminate luminance over nine orders of magnitude despite only three orders of contrast being available at a given moment. We would expect such a solution to be used for evaluating RPEs, which likewise have a very broad range. As such, this moderator analysis suggests that outcomes are not scaled to the range of magnitudes available on a *trial*, but the wider context of the *experiment*. Nevertheless, this is an interesting result, since it suggests that the expected value term against which RPEs are calculated may not simply be inherited from the midbrain dopamine system, or at least those midbrain dopaminergic neurons that have shown strong scaling effects (Tobler et al., 2005), and is thus relevant to the ongoing question of the afferents of the FRN.

5.5.2 Applications of the Present Findings

The FRN is a robustly elicited component, easy to study in human participants, and appears to encode an RPE. It may thus contribute to the daunting task of uncovering the network of neural events that give rise to subjective valuation by humans. Holroyd and Coles’ theory of the FRN was focused on its role in reinforcement learning, rather than its role as a general index of subjective value. The relationship between reinforcement learning and valuation is close however. The information concerning action-reward contingencies that is held in a reinforcement learning system presumably

strongly informs the valuation of the actions available to people in a given situation. Thus, if it can be measured (e.g. by the FRN) it has predictive power for human choice of the kind that neuroeconomics strives to attain.

The nature of the reinforcement learning system underlying the FRN is therefore pertinent. Reinforcement learning falls into two broad classes, model-free and model-based. Model-free reinforcement learning assigns values to actions based on the net reward they can expect to incur, without consideration of the actual outcomes that are produced. The values are updated in light of RPEs, but are termed “habit values” because they encode only the historical value of an action. Such learning is computationally efficient and information poor because the structure of rewards and the probabilities that follow an action is cached into a single value. Model-based reinforcement learning uses a model of the environment which represents actions, rewards, and intermediate states, and calculates values of actions by a tree search of this model. While more computationally expensive, this can be more quickly updated. A recent review of model-free and model-based reinforcement learning is provided by Walsh and Anderson (2014).

The relevance of this distinction to human choice is that model-based reinforcement learning is likely to be continuous with general cognition (Chater, 2009). Thus the degree to which choice on any one occasion is influenced by wider knowledge, by deliberative reasoning, or by verbal instruction will depend on the degree to which a model-free habitual system or a model-based belief system is dominant at that time. If the FRN can be established as belonging to one system or the other, it can be used as a much more direct means to investigate the relative contributions of habitual and belief based valuation to behavior, and assist in accounting for variations in both inter- and intra-individual choice that elude the revealed preferences method.

While the present demonstration that the FRN encodes an RPE places the debate on a much firmer footing, there has nevertheless been limited work on this important question. Hajcak et al. (2007) and Moser and Simons (2009) both showed a relationship of FRN amplitude to RPEs generated against subjective predictions but not to reinforcement history, implying the component might arise from model-based reinforcement learning, while Ichikawa, Siegle, Dombrowski, and Ohira (2010) found comparable contributions of subjective prediction and reinforcement history to FRN amplitude. However, Walsh and Anderson (2011b) found persuasive evidence against model-based reinforcement learning. They compared the FRN in cases where participants received verbal instruction on choice- outcome contingencies to cases where they did not. In the instruction condition, participants used this instruction, as shown by their behavior, thus adopting the given “model”. However, when unexpected outcomes, that is model-based RPEs, occurred, the FRN was initially insensitive to these. Its sensitivity developed only at the rate shown in the no-instruction condition suggesting it was dependent on a model-free history of reinforcement. A number of other authors have been able to show that FRN amplitude corresponds to the size of RPEs derived from a model-free reinforcement learning algorithm (Chase et al., 2011; Cohen & Ranganath, 2007; Philiastides, Biele, Vavatzanidis, Kazzner, & Heekeren, 2010). Other evidence for a model-free basis for the FRN comes from the demonstration that dopamine, the neurotransmitter implicated in generating the FRN, promotes model-free rather than model-based reinforcement learning (Wunderlich, Smittenaar, & Dolan, 2012). On current balance the evidence favours the FRN’s role in model-free reinforcement learning.

Insofar as model-free reinforcement learning is computationally cheap, it might be expected to occur by default, and indeed, to continue to compute valuations and associated RPEs even when a superior model-based reinforcement learning system was

guiding behavior. Bayer and Glimcher (2005), for example, showed that midbrain dopaminergic neurons, which are believed to underlie the FRN, showed firing patterns consistent with a model-free RPE and continued to behave in this fashion even when their effect on behavior was weak. In the case of the FRN itself, the component has in some cases been shown to predict choice in a way that is consistent with reinforcement learning (Cohen & Ranganath, 2007; Van der Helden, Boksem, & Blom, 2010; Yasuda, Sato, Miyawaki, Kumano, & Kuboki, 2004), but in other cases it has not (Mars, De Bruijn, Hulstijn, Miltner, & Coles, 2004; Mas-Herrero & Marco-Pallarés, 2014; San Martín, Appelbaum, Pearson, Huettel, & Woldorff, 2013; Yeung & Sanfey, 2004). In particular, Chase et al. (2011) showed that in a reversal learning task, the nature of which would be expected to engage model-based reinforcement learning, an FRN was observed that was well described by model-free reinforcement learning but which nevertheless did not predict behavior, suggesting it was over-ridden by a model-based system. Findings such as these suggest that the FRN might be used to predict behavior in situations promoting relatively automatic, fast judgments, what has been described by dual process theories as System 1 (Kahneman, 2003). Such valuation has been under-represented by the traditional methods of behavioral economics, which rely on stated (rather than observed) preferences in one-shot (rather than repeated) choices, which place prominence on deliberative processing. Perhaps the most serious challenge that the studies cited above pose for behavioral and neoclassical economics however, lies in the possibility that rather than value being constructed from multiple terms, as is suggested for example by prospect theory, quite separate independent valuations might be constructed which have differential access to behavior depending on circumstances.

Even while the precise nature of the valuation associated with the FRN remains unresolved, it may nevertheless serve as a biomarker for subjective value. It has been proposed in this regard for a range of psychopathologies such as hypomania and

depression (Bress, Smith, Foti, Klein, & Hajcak, 2012; Mason, O'Sullivan, Bentall, & El-Deredy, 2012) and pathological gambling (Hewig et al., 2010). Furthermore, recent studies have shown that variation in dopaminergic genes affects the component (Foti & Hajcak, 2012; Marco-Pallarés et al., 2009) raising the possibility that it might be used to investigate the proximate basis of genetic effects on behaviour. With the advent of mobile electroencephalography (EEG) setups that can be ready to use within minutes, the FRN may also provide a useful general measure of the subjective value of an outcome even in studies in which the brain is not the principal focus, much as other psychophysiological techniques such as skin conductance and pupillometry are used more broadly. As a dependent variable of subjective value it has a number of advantages over self-report. Asking participants to report on their valuations brings in extra processes which generally undermines the ecological validity of the study of “on-line” evaluation. Reported valuations may be participant to demand characteristics since participants are likely to be aware of at least some norms in economic preference, such as avoiding obvious inconsistencies and intransitivities. Self-report may also be affected by what reference point stated valuation is taken with respect to, which depends in turn on the framing of the question used to prompt self-report.

5.5.3 Implications for the Measurement of the RPE-FRN and FRN

We have distinguished between a response simply to valence, the FRN, well established in the literature, and a neural response to the valence and size of an RPE, the RPE-FRN, for which we have presented evidence here. The distinction is important for the testing of Holroyd and Coles’ theory. However it is not widely made in the literature and the comments below apply equally to both FRN and RPE-FRN

The present meta-analysis revealed a wide variation in methods used to quantify the FRN, and we have noted the role this may play both in failures of replication and inflation of false positives. We have also noted the variability of the waveforms

themselves. These two aspects are linked insofar as inconsistencies in FRN quantification possibly reflect the genuine attempt to best tailor analysis to a component of seemingly inconsistent character on an experiment-by-experiment basis. However, if, as we have argued, variability in the waveforms largely reflects the vagaries of component overlap rather than real variability in the FRN, then this latitude in quantification is harmful. For example, in the present meta-analysis, P2 and N2 peaks varied so much in their latency across experiments that while we initially intended to apply the GGA technique to a peak to peak measure, implemented in standardized intervals, we were unable to do so. This illustrates the point that whilst peaks might provide compelling landmarks by which to detect the FRN in any *individual* study, the lack of consistency *across* studies suggests the benefits of locking FRN quantification to simple waveform peaks may be illusory. The loose relationship between single waveform peaks and the underlying components has been cogently described by Luck (2005).

As such, measures based on difference waves are to be preferred. For the specific case of the RPE-FRN, a measurement interval of 270–300 ms is suggested by the present study since this captures the strongest effects of both magnitude and likelihood and is thus the best estimate of the RPE-FRN's latency. However the RPE-FRN in individual experiments may be participant to genuine latency differences and so, based on the course of the effect under both modulators, the interval 240–340 ms may be more appropriate. It should be noted that studies which more effectively decompose waveforms into constituent components, for example using principal components analysis, may reveal a rather different latency for the underlying RPE encoder, or encoders, as was shown in Chapter 4. Indeed Figure 33 suggests that such decomposition may well be necessary to fully isolate the individual components.

5.5.4 Evaluation of the GGA Technique

The GGA technique was developed because the great variety in how the FRN was quantified rendered conventional meta-analysis highly problematic. It is worthwhile assessing how this technique performed, partly in judging the present findings, but also for its future use in ERP meta-analysis. First, our concerns regarding the conventional meta-analysis of the FRN using effect sizes derived from diverse quantifications proved justified. When such a meta-analysis was performed it failed to find a significant effect of magnitude on the FRN, despite this effect being strongly present in an ideal conventional meta-analysis on original data. In contrast, the GGA technique was in close agreement. The conclusion we draw from the superior performance of the GGAs is that it is more important to employ an appropriate and consistent quantification of a component than to have access to the measures of within-study variance that use of published statistics provides. Of course the ideal meta-analysis achieved both of these objectives. However the GGA technique only requires access to published data. This has a great number of advantages. Most importantly, it avoids the large reduction in sample size that reliance on original authors inevitably entails. It substantially reduces the effort required to acquire data and convert it to a common format, and makes no demands at all on the original authors. It removes the uncertainty surrounding the number of studies that the meta-analysis will contain, allowing the viability of the exercise to be assessed in advance. It avoids the danger of bias arising from authors selectively complying with the request for original data depending on what they perceive the meta-analyzer's hypothesis to be. Finally, the technique can be used to guide the development of future work. If an effect, component, or other subset of the ERP in a published study was not selected for analysis within that study, there will not be any effect sizes on which to base a traditional meta-analysis. The GGA technique allows for post-hoc exploration of published ERPs, allowing the

researcher to approximate the effect sizes of previously disregarded data to guide the design of new empirical study, theory, or analysis technique.

The GGA technique has some disadvantages. Simple, rather than standardized effect sizes were used, meaning that the GGA meta-analysis could not down-weight studies with large variance, thus introducing some noise into hypothesis testing. The extent of this can be simply estimated from the validation dataset by calculating the correlation coefficient of the simple effect sizes of the GGA technique and the standardized effect sizes of the ideal conventional meta-analysis: the lower the correlation, the greater the noise introduced by failure to use standardized effect size. The value was $r = 0.8$, suggesting a moderate degree of noise introduced. This is, however, an overestimate of the problem insofar as standardized effect sizes themselves are not perfect because the standard deviations they are built from are themselves subject to sampling error. The remaining source of noise in the GGA technique consists in deviations between the digitized waveforms used for the meta-analysis and the original data, however comparisons with the validation dataset set suggest this is relatively small.

It should be noted that differences between experiments regarding the reference electrode, filters and baseline do not impact on the GGA technique since all contrasts and simple effect sizes are generated *within-experiment*, and so these extraneous factors can never become confounds for the simple generic reason that they are held constant at the point of generating simple effect sizes. It is true that FRN amplitude itself may be affected by these parameters, and that a poorly chosen reference electrode, for example, might reduce FRN amplitudes overall, concomitantly reduce effect sizes, and assist in rendering a meta-analysis non-significant. However in this regard GGAs do not differ from conventional meta-analysis. We simply note that because reference electrode is held constant within each study it does not confound the simple effect size generated for

that individual study, and since this meta-analysis is simply a collation of such simple effect sizes it likewise cannot be confounded by reference electrode.

We propose the GGA technique as a general method for meta-analysis of ERP components, not just the FRN. While inconsistency of measurement has been shown to be a particular problem for the FRN, this is also likely to be true to some degree of other components. Furthermore, even when conventional meta-analysis is applied, we still propose that this be performed in concert with a GGA analysis, to check there is no gross difference in results. As an accompanying method it also has the advantage that it allows the plotting of a waveform to accompany the reported effects. Individual ERP experiments ubiquitously plot an entire waveform despite their reported effects occurring in a small portion of the waveform since it provides a “sanity check” that the ERP shows a representative character and that the interval chosen for analysis is reasonable. A GGA waveform serves the same function in the case of a meta-analysis.

6 General Discussion

In the introduction to the thesis we raised some key questions regarding the FRN. In this section, we examine the extent to which the present study informs these questions.

6.1 The Reinforcement Learning Theory of the Error Related Negativity

The reinforcement learning theory of the error related negativity (RL-ERN) theory of Holroyd and Coles (2002) claims that an RPE carried from the midbrain dopamine system to the ACC produces an ERP, manifesting itself as the FRN when external feedback generates the RPE and as the ERN when internal monitoring generates the RPE. The present work supports some of these claims but not others. The claim that the FRN represents an RPE was bolstered by a demonstration in Chapters 3 and 4 that it could be distinguished from a salience response. Both experiments in those chapters extracted a frontocentral component, occurring in the conventional FRN interval of 200–350 ms which responded to +RPEs but not –RPEs, with an additional further encoding of just –RPEs in a later interval in the case of Pavlovian values. The claim that the FRN encodes an RPE was further bolstered by the demonstrations in Chapters 4 and 5, showing that the extracted FRN component was similar regardless of whether RPE size was manipulated by outcome likelihood or magnitude. This encourages the view that a valuation was performed, rather than a mere response to the psychological impact of unlikely or large events. Testing the claim that the FRN has its origins in the midbrain dopamine system is largely beyond the scope of this thesis, though we note that the asynchronous +RPE and –RPEs encoding responses shown in Chapter 3 are consistent with studies suggesting that midbrain dopamine neurons signal the size of +RPEs with an immediate phasic firing, and signal –RPEs by the duration of

firing inhibition. The claim that the FRN and the ERN arise from a common generator was strongly disconfirmed. The ERN occurs when a response is made that is not the response that has previously been rewarded. This event thus describes, specifically, a –RPE under instrumental conditioning. However, in Chapter 4 where an instrumental task was used, the FRN was not responsive to –RPEs, and indeed, no –RPE signal could be found. As such, the FRN and ERN appear to respond to +RPEs and –RPEs respectively and so must be functionally distinct. Finally, Holroyd and Coles’ claim that the FRN arises in the ACC did not receive support from this study which, after extraction of the FRN with PCA, found a source in the putamen.

6.2 Is the FRN a Response to +RPEs, –RPEs or Both?

The comparative negativity of the ERP waveform to –RPEs compared to +RPEs does not reveal whether this results from a negative shift in response to –RPEs or a positive shift to +RPEs or both. We critiqued previous attempts to answer this question which involved studying the simple effect of RPE size on +RPEs and –RPEs separately, implementing alternative methodologies in Chapters 3 and 4. In Chapter 3, a Pavlovian task revealed that, at the latency at which the bulk of FRN measurements are centred (~270 ms), the RPE response was generated by +RPEs alone, with a response to –RPEs shown in the later interval of 310–390 ms. In Chapter 4, an instrumental task also generated an FRN peaking at a latency of ~270 ms but decomposition with PCA revealed that this arose from the combination of an earlier asymmetrical salience response and a later (320–330 ms) bona fide RPE response to +RPEs only. This response was much stronger than that in the Pavlovian task in Chapter 3. Thus, the FRN *as typically measured*, appears to be a +RPE response. Furthermore, the absence of –RPE encoders in the decomposed ERPs suggests that there may be a general bias towards encoding of +RPEs over –RPEs in instrumental tasks. Clearly, this absence is at

odds with a key assumption of reinforcement learning theory: that an RPE encoder should be responsive to both the negative and positive portions of the reward prediction error continuum

6.3 Is the FRN Insensitive to Expected Value?

The question under consideration here concerns domain: are RPEs differently valued depending on whether the actual outcome is an absolute gain or loss. This is of interest in its own right but is also of methodological interest since, if so, domain may need to be controlled in order to adequately investigate RPEs. Chapter 2 revealed that the FRN, that is the difference wave of good and bad outcomes, was of greater amplitude for gain domain gambles. This contrasted markedly with participants' affective ratings in which they claimed a greater sensitivity to the outcome's valence in the loss domain. These ratings were in keeping with prospect theory, which is itself largely based on data gained from verbal report. As such, the findings suggest that prospect theoretic effects do not have their origin in the basic reinforcement learning apparatus on which the FRN is assumed to be based. The reasons for the increased sensitivity of the FRN towards the gain domain remain unclear. It is possible that this simply reflects a relatively high level orientation that participants took towards the task, viewing it as an attempt to maximise gains rather than minimise losses, and indeed a future avenue of research might be to attempt to manipulate such an orientation. Alternatively it is very possible that the constellation of +RPEs, go-tasks and dopamine that was discussed in Chapter 4 extends also to the gain domain insofar as when animals act to optimise outcomes they are more often than not doing so in order to obtain inherently desirable (gain domain) outcomes such as food.

Nevertheless, the domain effect in Chapter 2 appears to be dependent on the foregrounding of the domain variable in the task used. In Chapter 3, where participants

completed one long block of gambles in one domain followed by another block in the other domain, domain effects were not shown. Furthermore, moderator analysis in the meta-analysis revealed no difference in effect size when comparing gain domain experiments vs. mixed domain experiments (though the effect under analysis was the RPE-FRN, not the FRN). Provisionally we thus conclude that domain has not confounded RPE valence in the literature overall.

6.4 Is the FRN Indifferent to Whether RPE Size is Modulated by the Magnitude of the Outcome or its Likelihood?

While it has been claimed that the FRN is insensitive to reward magnitude, Chapters 3 and 4 presented clear evidence against this, as did the meta-analysis performed in Chapter 5. The question remains, however, whether or not the observed effects in these chapters arise from a single component responding to RPE size under both likelihood and magnitude manipulation. This is not easy to establish insofar as it requires a demonstration of the null hypothesis that the two modulators have no differential effect on the FRN. We are at present limited to saying that we have no reason to reject this null hypothesis. The FRN factor extracted by the PCAs of Chapter 4 was qualitatively extremely similar under both magnitude and likelihood modulation. The meta-analysis in Chapter 5 failed to find a significant moderating influence of modulator. It is true however that this moderator analysis was burdened by a confounding of modulator with the degree of control participants had over outcomes (tending to be much higher in likelihood experiments), thus greatly reducing the ANCOVA's power to extract a modulator effect once the control over outcome effect had been demonstrated. As such we consider the question still to be open.

6.5 Where and when does the FRN occur?

One of the contributions of the present work is to attempt to improve the precision with which the FRN and related components are extracted. Chapters Two, Three and Four represent a progression in this regard, from a conventional measurement of the FRN in a fixed interval, to a data-led temporal parsing of the waveform, to a further parsing based on individual and topographical differences using PCA

This proves to be a somewhat misconceived question insofar as the present work suggests that activity that has been historically attributed to the FRN may be attributable to multiple components, not all of which code RPEs. The meta-analysis in Chapter 5 (see Figure 33) suggests the presence of multiple components responding to some combination of valence, RPE size and their interaction. Additionally, Chapter 4 showed how an FRN observed at a latency of ~270 ms in grand average waveforms proved, after decomposition by PCA, to be the sum of two quite separate effects at earlier and later latencies. Conversely we have shown, that RPE encoding may occur at latencies that lie beyond the interval in which the FRN is typically measured. This was shown by the late –RPE encoding shown in Chapter 3.

The meta-analysis in Chapter 5 provides an unusually robust basis for determining the latency of the FRN, as typically operationalised as a simple difference wave of good and bad outcomes. This difference wave occupied a wide interval running from 150–400 ms. However, this is likely to be contaminated by P3 effects at the later latencies, and should therefore be restricted to the interval in which the signal was greater at frontocentral than parietal sites. Since this is not possible with the data available to the meta-analysis, the most useful landmark is the latency of the FRN's peak. This occurred at 276 ms.

We have made the case for distinguishing between this valence effect (the FRN) and an effect whereby valence is modulated by RPE size (the RPE-FRN), since this

latter effect is characteristic of an RPE encoder. The peak amplitude of this component was at 283 ms (after collapsing together the two modulators).

While their findings are less robust insofar as they are based on single experiments, Chapters 3 and 4 are also relevant to identifying the latency of the FRN. Both chapters suggested the presence of a frontal salience coding component, dubbed the P2, that immediately preceded the FRN's peak and therefore overlapped its earlier portion. Ensuring that FRN measurements are not contaminated by this salience component is therefore an important consideration. In principle, contamination should not be an issue insofar as the FRN is extracted by a contrast of +RPEs and -RPEs, while salience is extracted by a contrast of small and large RPEs: thus the components are formally non-confounded. Nevertheless, Chapter 4 revealed that the extracted P2 had an asymmetrical salience response, producing a stronger response to +RPEs than -RPEs. As such, the earlier portion of an FRN difference wave may not arise from the FRN at all. The situation is unclear however, because while it is tempting to view the extracted factors in Chapter 4 as a reification of the real underlying neural generators, these factors are nevertheless still descriptions of scalp recorded activity and so retain at least some effects of component overlap. Thus a subsequent experiment which was subjected to PCA might well re-allocate some of the variance attributed here to the P2 to the following FRN, perhaps removing the observed asymmetry in the P2's salience response. Alternatively, this effect might be achieved by the extraction of an entirely new, early +RPE encoding factor that inherited the asymmetrical portion of the salience response.

The question of the location of the neural generator of the FRN was not the central subject of this thesis. Nevertheless we performed source localisation in chapter 4. The signal undergoing localisation was unusually specific insofar as it was not only an individual factor extracted by a preceding PCA, but also specifically a factor sensitive to

the size and valence of an RPE. This factor localised to the putamen in keeping with a number of recent studies on PCA decomposed FRN data.

6.6 Is the RPE Indexed by the FRN a General Valuation or Specific to a Reinforcement Learning Module?

As we note above, this question presumes the existence of a single component where we have presented evidence for many. As such we can restate the question: what kinds of valuation can be observed in feedback-locked ERPs? Balleine et al. (2008) have proposed three forms of valuation in reinforcement learning: Pavlovian values, and two instrumental values: habit and goal values. RPEs generated with respect to Pavlovian values were shown in Chapter 3. These signals were small however, compared to the instrumental +RPE response shown in Chapter 4. The small size of RPEs under Pavlovian tasks was also shown by the moderator analysis of control over outcome in Chapter 5.

Broadly conceived, the FRN appears to be an RPE constructed against a value derived from instrumental learning. Though the experiment was not devised to distinguish goal from habit values, some circumstantial evidence from Chapter 2 suggests the FRN is likely to arise from an RPE constructed against a habit value. This is because goal values are assumed to be informed by a wider model of the environment than a simple reinforcement schedule and this chapter showed marked differences between the behavior of the FRN and verbal affective ratings. Overall, Holroyd and Coles' characterisation of the FRN as a simple reinforcement learning signal, linking a midbrain prediction error to action selection was supported, albeit for +RPEs only.

6.7 How should the FRN be studied?

We end the thesis with a few methodological recommendations for the design and analysis of FRN experiments.

We propose that the independent variable (typically RPE size or sign, or the combination of these) be manipulated as a continuous variable wherever possible because this increases power and improves the chance of isolating an RPE encoder from other components. In this regard, one virtue of modelling RPEs using a Rescorla-Wagner algorithm is that even stimuli which are naturally categorical can be converted to an RPE which is a continuous variable. Nevertheless, even though categorical stimuli can be converted to a continuous independent variable in this fashion, categorical stimuli should still be avoided in favour of stimuli that themselves vary along some continuous dimension. This is because such continuous stimuli undermine the generation of components devoted to detecting a mismatch between a feedback stimulus and some internal representation corresponding to either a hoped-for, or feared, stimulus. Such matching effects have been shown to be a key confound by Jia et al. (2007) and Donkers et al. (2005). The confound is particularly insidious in that there is no way of controlling such internal representations, and yet they may nevertheless be influenced by subtle aspects of tasks. Since FRN experiments based on likelihood modulation are much more likely to involve a binary stimulus set, this suggests it may be preferable to modulate the FRN using magnitude.

(Banis & Lorist, 2012; Bellebaum, Kobza, Thiele, & Daum, 2010, 2011; Franken, Van den Berg, & Van Strien, 2010; Hajcak, Holroyd, Moser, & Simons, 2005; HajiHosseini, Rodriguez-Fornells, & Marco-Pallares, 2012; Holroyd, Krigolson, Baker, Lee, & Gibson, 2009; Holroyd, Nieuwenhuis, Yeung, & Cohen, 2003; Kamarajan et al., 2009; Larson, Kelly, Stigge-Kaufman, Schmalfuss, & Perlstein, 2007; Liao, Gramann, Feng, Deák, & Li, 2011; Luo & Qu, 2013; Morris, Holroyd, Mann-Wrobel, & Gold, 2011; Mushtaq, M., Mon-Williams, & Schaefer, unpublished data ; Nittono, Otsuka, & Ullsperger, 2008; Onoda, Abe, & Yamaguchi, 2010;

Roberts, Infantolino, Stanley, & Simons, unpublished data; T. D. Sambrook & Goslin, unpublished data experiment 1, unpublished data experiment 2, unpublished data experiment 3; San Martin et al., 2010; Santesso, Dzyundzyak, & Segalowitz, 2011; Sato et al., 2005; Schuermann, Endrass, & Kathmann, 2012; Toyomaki & Murohashi, 2005; Van den Berg, Franken, & Muris, 2011; Walsh & Anderson, unpublished data experiment 1, unpublished data experiment 2; Walsh & Anderson, 2011a, 2013; Yi et al., 2012; Yu & Zhou, 2006, 2009; Zottoli & Grose-Fifer, 2012)

7 References

- Asaad, W. F., & Eskandar, E. N. (2011). Encoding of Both Positive and Negative Reward Prediction Errors by Neurons of the Primate Lateral Prefrontal Cortex and Caudate Nucleus. *Journal of Neuroscience*, 31(49), 17772-17787. doi: Doi 10.1523/Jneurosci.3793-11.2011
- Baguley, T. (2009). Standardized or simple effect size: What should be reported? *British Journal of Psychology*, 100, 603-617. doi: Doi 10.1348/000712608x377117
- Balleine, B. W., Daw, N. D., & O'Doherty, J. P. (2008). Multiple forms of value learning and the function of dopamine. *Neuroeconomics: Decision making and the brain*, 367-385.
- Banis, S., & Lorist, M. M. (2012). Acute noise stress impairs feedback processing. *Biological Psychology*, 91(2), 163-171. doi: DOI 10.1016/j.biopsycho.2012.06.009
- Bartra, O., McGuire, J. T., & Kable, J. W. (2013). The valuation system: a coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. *Neuroimage*, 76, 412-427. doi: <http://dx.doi.org/10.1016/j.neuroimage.2013.02.063>

- Bayer, H. M., & Glimcher, P. W. (2005). Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron*, 47(1), 129-141. doi: DOI 10.1016/j.neuron.2005.05.020
- Bayer, H. M., Lau, B., & Glimcher, P. W. (2007). Statistics of midbrain dopamine neuron spike trains in the awake primate. *Journal of Neurophysiology*, 98(3), 1428-1439. doi: DOI 10.1152/jn.01140.2006
- Bellebaum, C., & Daum, I. (2008). Learning-related changes in reward expectancy are reflected in the feedback-related negativity. *European Journal of Neuroscience*, 27(7), 1823-1835. doi: DOI 10.1111/j.1460-9568.2008.06138.x
- Bellebaum, C., Kobza, S., Thiele, S., & Daum, I. (2010). It Was Not MY Fault: Event-Related Brain Potentials in Active and Observational Learning from Feedback. *Cerebral Cortex*, 20(12), 2874-2883. doi: DOI 10.1093/cercor/bhq038
- Bellebaum, C., Kobza, S., Thiele, S., & Daum, I. (2011). Processing of Expected and Unexpected Monetary Performance Outcomes in Healthy Older Subjects. *Behavioral Neuroscience*, 125(2), 241-251. doi: Doi 10.1037/A0022536
- Bellebaum, C., Polezzi, D., & Daum, I. (2010). It is less than you expected: The feedback-related negativity reflects violations of reward magnitude expectations. *Neuropsychologia*, 48(11), 3343-3350. doi: DOI 10.1016/j.neuropsychologia.2010.07.023
- Bond, C. F., Wiitala, W. L., & Richard, F. D. (2003). Meta-analysis of raw mean differences. *Psychological Methods*, 8(4), 406-418. doi: Doi 10.1037/1082-989x.8.4.406
- Boureau, Y. L., & Dayan, P. (2011). Opponency Revisited: Competition and Cooperation Between Dopamine and Serotonin. *Neuropsychopharmacology*, 36(1), 74-97. doi: Doi 10.1038/Npp.2010.151

- Bress, J. N., Smith, E., Foti, D., Klein, D. N., & Hajcak, G. (2012). Neural response to reward and depressive symptoms in late childhood to early adolescence. *Biological Psychology*, 89(1), 156-162.
- Bromberg-Martin, E. S., Matsumoto, M., & Hikosaka, O. (2010). Dopamine in Motivational Control: Rewarding, Aversive, and Alerting. *Neuron*, 68(5), 815-834. doi: DOI 10.1016/j.neuron.2010.11.022
- Brown, P., & Molliver, M. E. (2000). Dual serotonin (5-HT) projections to the nucleus accumbens core and shell: Relation of the 5-HT transporter to amphetamine-induced neurotoxicity. *Journal of Neuroscience*, 20(5), 1952-1963.
- Brunia, C., van Boxtel, G., & Bocker, K. (2011). Negative slow waves as indices of anticipation: the Bereitschaftspotential, the contingent negative variation, and the stimulus-preceding negativity. *The Oxford handbook of event-related potential components*.
- Bunzeck, N., Dayan, P., Dolan, R. J., & Duzel, E. (2010). A Common Mechanism for Adaptive Scaling of Reward and Novelty. *Human Brain Mapping*, 31(9), 1380-1394. doi: Doi 10.1002/Hbm.20939
- Campbell-Meiklejohn, D., Wakeley, J., Herbert, V., Cook, J., Scollo, P., Ray, M. K., . . . Rogers, R. D. (2011). Serotonin and Dopamine Play Complementary Roles in Gambling to Recover Losses. *Neuropsychopharmacology*, 36(2), 402-410. doi: Doi 10.1038/Npp.2010.170
- Caplin, A., & Dean, M. (2008). Axiomatic methods, dopamine and reward prediction error. *Current Opinion in Neurobiology*, 18(2), 197-202. doi: DOI 10.1016/j.conb.2008.07.007
- Carlson, J. M., Foti, D., Mujica-Parodi, L. R., Harmon-Jones, E., & Hajcak, G. (2011). Ventral striatal and medial prefrontal BOLD activation is correlated with

- reward-related electrocortical activity: A combined ERP and fMRI study.
Neuroimage, 57(4), 1608-1616. doi: DOI 10.1016/j.neuroimage.2011.05.037
- Chase, H. W., Swainson, R., Durham, L., Benham, L., & Cools, R. (2011). Feedback-related Negativity Codes Prediction Error but Not Behavioral Adjustment during Probabilistic Reversal Learning. *Journal of Cognitive Neuroscience*, 23(4), 936-946. doi: DOI 10.1162/jocn.2010.21456
- Chater, N. (2009). Rational and mechanistic perspectives on reinforcement learning. *Cognition*, 113(3), 350-364. doi: DOI 10.1016/j.cognition.2008.06.014
- Cohen, J. (1983). The cost of dichotomization. *Applied Psychological Measurement*, 7, 249-253.
- Cohen, M. X., Cavanagh, J. F., & Slagter, H. A. (2011). Event-Related Potential Activity in the Basal Ganglia Differentiates Rewards From Nonrewards: Temporospacial Principal Components Analysis and Source Localization of the Feedback Negativity: Commentary. *Human Brain Mapping*, 32(12), 2270-2271. doi: Doi 10.1002/Hbm.21358
- Cohen, M. X., Elger, C. E., & Ranganath, C. (2007). Reward expectation modulates feedback-related negativity and EEG spectra. *Neuroimage*, 35(2), 968-978. doi: DOI 10.1016/j.neuroimage.2006.11.056
- Cohen, M. X., & Ranganath, C. (2007). Reinforcement learning signals predict future decisions. *Journal of Neuroscience*, 27(2), 371-378. doi: Doi 10.1523/Jneurosci.4421-06.2007
- Daw, N. D., Kakade, S., & Dayan, P. (2002). Opponent interactions between serotonin and dopamine. *Neural Networks*, 15(4-6), 603-616.
- De Martino, B., Kumaran, D., Holt, B., & Dolan, R. J. (2009). The Neurobiology of Reference-Dependent Value Computation. *Journal of Neuroscience*, 29(12), 3833-3842. doi: Doi 10.1523/Jneurosci.4832-08.2009

- De Martino, B., Kumaran, D., Seymour, B., & Dolan, R. J. (2006). Frames, biases, and rational decision-making in the human brain. *Science*, 313(5787), 684-687. doi: DOI 10.1126/science.1128356
- Dien, J. (2010a). The ERP PCA Toolkit: An open source program for advanced statistical analysis of event-related potential data. *Journal of Neuroscience Methods*, 187(1), 138-145. doi: DOI 10.1016/j.jneumeth.2009.12.009
- Dien, J. (2010b). Evaluating two-step PCA of ERP data with Geomin, Infomax, Oblimin, Promax, and Varimax rotations. *Psychophysiology*, 47(1), 170-183. doi: DOI 10.1111/j.1469-8986.2009.00885.x
- Dien, J., Beal, D. J., & Berg, P. (2005). Optimizing principal components analysis of event-related potentials: Matrix type, factor loading weighting, extraction, and rotations. *Clinical Neurophysiology*, 116(8), 1808-1825. doi: DOI 10.1016/j.clinph.2004.11.025
- Dien, J., Khoe, W., & Mangun, G. R. (2007). Evaluation of PCA and ICA of simulated ERPs: Promax vs. infomax rotations. *Human Brain Mapping*, 28(8), 742-763. doi: Doi 10.1002/Hbm.20304
- Dien, J., Spencer, K. M., & Donchin, E. (2003). Localization of the event-related potential novelty response as defined by principal components analysis. *Cognitive Brain Research*, 17(3), 637-650. doi: Doi 10.1016/S0926-6410(03)00188-5
- Donkers, F. C. L., Nieuwenhuis, S., & van Boxtel, G. J. M. (2005). Mediofrontal negativities in the absence of responding. *Cognitive Brain Research*, 25(3), 777-787. doi: DOI 10.1016/j.cogbrainres.2005.09.007
- Donkers, F. C. L., & van Boxtel, G. J. M. (2005). Mediofrontal negativities to averted gains and losses in the slot-machine task - A further investigation. *Journal of Psychophysiology*, 19(4), 256-262. doi: Doi 10.1027/0269-8803.19.4.256

- Dunlap, W. P., Cortina, J. M., Vaslow, J. B., & Burke, M. J. (1996). Meta-analysis of experiments with matched groups or repeated measures designs. *Psychological Methods*, 1(2), 170-177. doi: Doi 10.1037/1082-989x.1.2.170
- Duval, S., & Tweedie, R. (2000). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, 56(2), 455-463. doi: DOI 10.1111/j.0006-341X.2000.00455.x
- Eppinger, B., Mock, B., & Kray, J. (2009). Developmental differences in learning and error processing: Evidence from ERPs. *Psychophysiology*, 46(5), 1043-1053. doi: DOI 10.1111/j.1469-8986.2009.00838.x
- Esber, G. R., & Haselgrove, M. (2011). Reconciling the influence of predictiveness and uncertainty on stimulus salience: a model of attention in associative learning. *Proceedings of the Royal Society B-Biological Sciences*, 278(1718), 2553-2561. doi: DOI 10.1098/rspb.2011.0836
- Field, A. P., & Gillett, R. (2010). How to do a meta-analysis. *British Journal of Mathematical & Statistical Psychology*, 63(3), 665-694. doi: Doi 10.1348/000711010x502733
- Folstein, J. R., & Van Petten, C. (2008). Influence of cognitive control and mismatch on the N2 component of the ERP: A review. *Psychophysiology*, 45(1), 152-170. doi: DOI 10.1111/j.1469-8986.2007.00602.x
- Foti, D., & Hajcak, G. (2012). Genetic variation in dopamine moderates neural response during reward anticipation and delivery: Evidence from event - related potentials. *Psychophysiology*, 49(5), 617-626.
- Foti, D., Weinberg, A., Dien, J., & Hajcak, G. (2011a). Event-Related Potential Activity in the Basal Ganglia Differentiates Rewards From Nonrewards: Response to Commentary. *Human Brain Mapping*, 32(12), 2267-2269. doi: Doi 10.1002/Hbm.21357

- Foti, D., Weinberg, A., Dien, J., & Hajcak, G. (2011b). Event-Related Potential Activity in the Basal Ganglia Differentiates Rewards from Nonrewards: Temporospatial Principal Components Analysis and Source Localization of the Feedback Negativity. *Human Brain Mapping, 32*(12), 2207-2216. doi: Doi 10.1002/Hbm.21182
- Franken, I. H., Van den Berg, I., & Van Strien, J. W. (2010). Individual differences in alcohol drinking frequency are associated with electrophysiological responses to unexpected nonrewards. *Alcoholism: Clinical and Experimental Research, 34*(4), 702-707.
- Garrison, J., Erdeniz, B., & Done, J. (2013). Prediction error in reinforcement learning: A meta-analysis of neuroimaging studies. *Neuroscience and Biobehavioral Reviews, 37*(7), 1297-1310. doi: DOI 10.1016/j.neubiorev.2013.03.023
- Gehring, W. J., & Willoughby, A. R. (2002). The medial frontal cortex and the rapid processing of monetary gains and losses. *Science, 295*(5563), 2279-2282.
- Gu, R. L., Lei, Z. H., Broster, L., Wu, T. T., Jiang, Y., & Luo, Y. J. (2011). Beyond valence and magnitude: A flexible evaluative coding system in the brain. *Neuropsychologia, 49*(14), 3891-3897. doi: DOI 10.1016/j.neuropsychologia.2011.10.006
- Hajcak, G., Holroyd, C. B., Moser, J. S., & Simons, R. F. (2005). Brain potentials associated with expected and unexpected good and bad outcomes. *Psychophysiology, 42*(2), 161-170. doi: DOI 10.1111/j.1469-8986.2005.00278.x
- Hajcak, G., Moser, J. S., Holroyd, C. B., & Simons, R. F. (2006). The feedback-related negativity reflects the binary evaluation of good versus bad outcomes. *Biological Psychology, 71*(2), 148-154. doi: DOI 10.1016/j.biopsycho.2005.04.001

- Hajcak, G., Moser, J. S., Holroyd, C. B., & Simons, R. F. (2007). It's worse than you thought: The feedback negativity and violations of reward prediction in gambling tasks. *Psychophysiology*, *44*(6), 905-912. doi: DOI 10.1111/j.1469-8986.2007.00567.x
- HajiHosseini, A., Rodriguez-Fornells, A., & Marco-Pallares, J. (2012). The role of beta-gamma oscillations in unexpected rewards processing. *Neuroimage*, *60*(3), 1678-1685. doi: DOI 10.1016/j.neuroimage.2012.01.125
- Hauk, O., Davis, M. H., Ford, M., Pulvermuller, F., & Marslen-Wilson, W. D. (2006). The time course of visual word recognition as revealed by linear regression analysis of ERP data. *Neuroimage*, *30*(4), 1383-1400. doi: DOI 10.1016/j.neuroimage.2005.11.048
- Hauser, T. U., Iannaccone, R., Stämpfli, P., Drechsler, R., Brandeis, D., Walitza, S., & Brem, S. (2014). The feedback-related negativity (FRN) revisited: New insights into the localization, meaning and network organization. *Neuroimage*, *84*, 159-168.
- Heidbreder, C. A., Hedou, G., & Feldon, J. (1999). Behavioral neurochemistry reveals a new functional dichotomy in the shell subregion of the nucleus accumbens. *Progress in Neuro-Psychopharmacology & Biological Psychiatry*, *23*(1), 99-132. doi: Doi 10.1016/S0278-5846(98)00094-3
- Hewig, J., Kretschmer, N., Trippe, R. H., Hecht, H., Coles, M. G., Holroyd, C. B., & Miltner, W. H. (2010). Hypersensitivity to reward in problem gamblers. *Biological Psychiatry*, *67*(8), 781-783.
- Hollerman, J. R., & Schultz, W. (1998). Dopamine neurons report an error in the temporal prediction of reward during learning. *Nature Neuroscience*, *1*(4), 304-309.

- Holroyd, C. B., & Coles, M. G. H. (2002). The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. *Psychological Review*, 109(4), 679-709. doi: Doi 10.1037//0033-295x.109.4.679
- Holroyd, C. B., Krigolson, O. E., Baker, R., Lee, S., & Gibson, J. (2009). When is an error not a prediction error? An electrophysiological investigation. *Cognitive Affective & Behavioral Neuroscience*, 9(1), 59-70. doi: Doi 10.3758/Cabn.9.1.59
- Holroyd, C. B., Krigolson, O. E., & Lee, S. (2011). Reward positivity elicited by predictive cues. *Neuroreport*, 22(5), 249-252. doi: Doi 10.1097/Wnr.0b013e328345441d
- Holroyd, C. B., Larsen, J. T., & Cohen, J. D. (2004). Context dependence of the event-related brain potential associated with reward and punishment. *Psychophysiology*, 41(2), 245-253. doi: DOI 10.1111/j.1469-8986.2004.00152.x
- Holroyd, C. B., Nieuwenhuis, S., Yeung, N., & Cohen, J. D. (2003). Errors in reward prediction are reflected in the event-related brain potential. *Neuroreport*, 14(18), 2481-2484. doi: DOI 10.1097/01.wnr.0000099601/41403.a5
- Holroyd, C. B., Pakzad-Vaezi, K. L., & Krigolson, O. E. (2008). The feedback correct-related positivity: Sensitivity of the event-related brain potential to unexpected positive feedback. *Psychophysiology*, 45(5), 688-697. doi: DOI 10.1111/j.1469-8986.2008.00668.x
- Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, 30(2), 179-185.
- Hunter, J. E., & Schmidt, F. L. (2004). *Methods of meta-analysis: Correcting error and bias in research findings*: Sage.
- Huys, Q. J. M., Cools, R., Golzer, M., Friedel, E., Heinz, A., Dolan, R. J., & Dayan, P. (2011). Disentangling the Roles of Approach, Activation and Valence in

Instrumental and Pavlovian Responding. *Plos Computational Biology*, 7(4). doi: ARTN e1002028 DOI 10.1371/journal.pcbi.1002028

Ichikawa, N., Siegle, G. J., Dombrovski, A., & Ohira, H. (2010). Subjective and model-estimated reward prediction: Association with the feedback-related negativity (FRN) and reward prediction error in a reinforcement learning task.

International Journal of Psychophysiology, 78(3), 273-283. doi: DOI 10.1016/j.ijpsycho.2010.09.001

Jia, S. W., Li, H., Luo, Y. J., Chen, A. T., Wang, B. X., & Zhou, X. L. (2007).

Detecting perceptual conflict by the feedback-related negativity in brain potentials. *Neuroreport*, 18(13), 1385-1388. doi: Doi

10.1097/Wnr.0b013e3282c48a90

Joshua, M., Adler, A., Mitelman, R., Vaadia, E., & Bergman, H. (2008). Midbrain

Dopaminergic Neurons and Striatal Cholinergic Interneurons Encode the Difference between Reward and Aversive Events at Different Epochs of

Probabilistic Classical Conditioning Trials. *Journal of Neuroscience*, 28(45), 11673-11684. doi: Doi 10.1523/Jneurosci.3839-08.2008

Kahneman, D. (2003). Maps of bounded rationality: Psychology for behavioral

economics. *American Economic Review*, 93(5), 1449-1475. doi: Doi

10.1257/000282803322655392

Kahneman, D., & Tversky, A. (1979). Prospect Theory - Analysis of Decision under

Risk. *Econometrica*, 47(2), 263-291. doi: Doi 10.2307/1914185

Kahnt, T., Park, S. Q., Cohen, M. X., Beck, A., Heinz, A., & Wrase, J. (2009). Dorsal

Striatal-Midbrain Connectivity in Humans Predicts How Reinforcements Are

Used to Guide Decisions. *Journal of Cognitive Neuroscience*, 21(7), 1332-1345.

doi: DOI 10.1162/jocn.2009.21092

- Kamarajan, C., Porjesz, B., Rangaswamy, M., Tang, Y. Q., Chorlian, D. B., Padmanabhapillai, A., . . . Begleiter, H. (2009). Brain signatures of monetary loss and gain: Outcome-related potentials in a single outcome gambling task. *Behavioural Brain Research*, 197(1), 62-76. doi: DOI 10.1016/j.bbr.2008.08.011
- Kennerley, S. W., Dahmubed, A. F., Lara, A. H., & Wallis, J. D. (2009). Neurons in the Frontal Lobe Encode the Value of Multiple Decision Variables. *Journal of Cognitive Neuroscience*, 21(6), 1162-1178. doi: DOI 10.1162/jocn.2009.21100
- Knutson, B., Adams, C. M., Fong, G. W., & Hommer, D. (2001). Anticipation of increasing monetary reward selectively recruits nucleus accumbens. *Journal of Neuroscience*, 21(16), art. no.-RC159.
- Knutson, B., Wimmer, G. E., Rick, S., Hollon, N. G., Prelec, D., & Loewenstein, G. (2008). Neural antecedents of the endowment effect. *Neuron*, 58(5), 814-822. doi: DOI 10.1016/j.neuron.2008.05.018
- Kobayashi, S., Nomoto, K., Watanabe, M., Hikosaka, O., Schultz, W., & Sakagami, M. (2006). Influences of rewarding and aversive outcomes on activity in macaque lateral prefrontal cortex. *Neuron*, 51(6), 861-870. doi: DOI 10.1016/j.neuron.2006.08.031
- Kreussel, L., Hewig, J., Kretschmer, N., Hecht, H., Coles, M. G. H., & Miltner, W. H. R. (2012). The influence of the magnitude, probability, and valence of potential wins and losses on the amplitude of the feedback negativity. *Psychophysiology*, 49(2), 207-219. doi: DOI 10.1111/j.1469-8986.2011.01291.x
- Langer, E. J. (1975). The illusion of control. *Journal of personality and social psychology*, 32(2), 311.
- Larson, M. J., Kelly, K. G., Stigge-Kaufman, D. A., Schmalfluss, I. M., & Perlstein, W. M. (2007). Reward context sensitivity impairment following severe TBI: An event-related potential investigation. *Journal of the International*

Neuropsychological Society, 13(4), 615-625. doi: Doi

10.1017/S1355617707070762

Liao, Y., Gramann, K., Feng, W., Deák, G. O., & Li, H. (2011). This ought to be good:

Brain activity accompanying positive and negative expectations and outcomes.

Psychophysiology, 48(10), 1412-1419. doi: [http://dx.doi.org/10.1111/j.1469-](http://dx.doi.org/10.1111/j.1469-8986.2011.01205.x)

8986.2011.01205.x

Luck, S. J. (2005). *An introduction to the event-related potential technique*. Cambridge,

Mass.: MIT Press.

Luo, Q. L., & Qu, C. (2013). Comparison Enhances Size Sensitivity: Neural Correlates

of Outcome Magnitude Processing. *Plos One*, 8(8). doi: ARTN e71186 DOI

10.1371/journal.pone.0071186

Marco-Pallarés, J., Cucurell, D., Cunillera, T., Krämer, U. M., Càmarà, E., Nager,

W., . . . Münte, T. F. (2009). Genetic Variability in the Dopamine System

(Dopamine Receptor D4, Catechol-< i> O</i>-Methyltransferase) Modulates

Neurophysiological Responses to Gains and Losses. *Biological Psychiatry*,

66(2), 154-161.

Marco-Pallares, J., Cucurell, D., Munte, T. F., Strien, N., & Rodriguez-Fornells, A.

(2011). On the number of trials needed for a stable feedback-related negativity.

Psychophysiology, 48(6), 852-860. doi: DOI 10.1111/j.1469-8986.2010.01152.x

Marco-Pallares, J., Kramer, U. M., Strehl, S., Schroder, A., & Munte, T. F. (2010).

When decisions of others matter to me: an electrophysiological analysis. *Bmc*

Neuroscience, 11. doi: Artn 86 Doi 10.1186/1471-2202-11-86

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and

MEG-data. *Journal of Neuroscience Methods*, 164(1), 177-190. doi: DOI

10.1016/j.jneumeth.2007.03.024

Mars, R. B., De Bruijn, E. R., Hulstijn, W., Miltner, W. H., & Coles, M. G. (2004).

What if I told you:“You were wrong”? Brain potentials and behavioral adjustments elicited by feedback in a time-estimation task. *Errors, conflicts, and the brain. Current opinions on performance monitoring*, 129-134.

Martin, L. E., Potts, G. F., Burton, P. C., & Montague, P. R. (2009).

Electrophysiological and hemodynamic responses to reward prediction violation. *Neuroreport*, 20(13), 1140-1143. doi: Doi 10.1097/Wnr.0b013e32832f0dca

Mas-Herrero, E., & Marco-Pallarés, J. (2014). Frontal theta oscillatory activity is a

common mechanism for the computation of unexpected outcomes and learning rate. *Journal of Cognitive Neuroscience*, 26(3), 447-458.

Mason, L., O'Sullivan, N., Bentall, R. P., & El-Deredy, W. (2012). Better Than I

Thought: Positive Evaluation Bias in Hypomania. *Plos One*, 7(10). doi: ARTN e47754 DOI 10.1371/journal.pone.0047754

Matsumoto, M., Matsumoto, K., Abe, H., & Tanaka, K. (2007). Medial prefrontal cell

activity signaling prediction errors of action values. *Nature Neuroscience*, 10(5), 647-656. doi: Doi 10.1038/Nn1890

Metereau, E., & Dreher, J. C. (2013). Cerebral Correlates of Salient Prediction Error for

Different Rewards and Punishments. *Cerebral Cortex*, 23(2), 477-487. doi: DOI 10.1093/cercor/bhs037

Mileykovskiy, B., & Morales, M. (2011). Duration of Inhibition of Ventral Tegmental

Area Dopamine Neurons Encodes a Level of Conditioned Fear. *Journal of Neuroscience*, 31(20), 7471-7476. doi: Doi 10.1523/Jneurosci.5731-10.2011

Miller, R. R., Barnet, R. C., & Grahame, N. J. (1995). Assessment of the Rescorla-

Wagner Model. *Psychological Bulletin*, 117(3), 363-386. doi: Doi 10.1037/0033-2909.117.3.363

- Miltner, W. H. R., Braun, C. H., & Coles, M. G. H. (1997). Event-related brain potentials following incorrect feedback in a time-estimation task: Evidence for a "generic" neural system for error detection. *Journal of Cognitive Neuroscience*, 9(6), 788-798. doi: DOI 10.1162/jocn.1997.9.6.788
- Monosov, I. E., & Hikosaka, O. (2012). Regionally Distinct Processing of Rewards and Punishments by the Primate Ventromedial Prefrontal Cortex. *Journal of Neuroscience*, 32(30), 10318-10330. doi: Doi 10.1523/Jneurosci.1801-12.2012
- Morris, S. E., Holroyd, C. B., Mann-Wrobel, M. C., & Gold, J. M. (2011). Dissociation of response and feedback negativity in schizophrenia: electrophysiological and computational evidence for a deficit in the representation of value. *Frontiers in Human Neuroscience*, 5. doi: Artn 123 Doi 10.3389/Fnhum.2011.00123
- Moser, J. S., & Simons, R. F. (2009). The neural consequences of flip-flopping: The feedback-related negativity and salience of reward prediction. *Psychophysiology*, 46(2), 313-320. doi: DOI 10.1111/j.1469-8986.2008.00760.x
- Mushtaq, F., M., W. R., Mon-Williams, M., & Schaefer, A. (unpublished data).
University of Leeds & Monash University.
- Nieuwenhuis, S., Aston-Jones, G., & Cohen, J. D. (2005). Decision making, the p3, and the locus coeruleus-norepinephrine system. *Psychological Bulletin*, 131(4), 510-532. doi: Doi 10.1037/0033-2909.131.4.510
- Nittono, H., Otsuka, Y., & Ullsperger, P. (2008). Asymmetrical effects of frequent gains and frequent losses in a gambling task. *Neuroreport*, 19(13), 1345-1349. doi: Doi 10.1097/Wnr.0b013e32830c21a8
- Oliveira, F. T. P., McDonald, J. J., & Goodman, D. (2007). Performance monitoring in the anterior Cingulate is not all error related: Expectancy deviation and the representation of action-outcome associations. *Journal of Cognitive Neuroscience*, 19(12), 1994-2004. doi: DOI 10.1162/jocn.2007.19.12.1994

- Onoda, K., Abe, S., & Yamaguchi, S. (2010). Feedback-related negativity is correlated with unplanned impulsivity. *Neuroreport*, 21(10), 736-739. doi: Doi 10.1097/Wnr.0b013e32833bfd36
- Padoa-Schioppa, C., & Assad, J. A. (2006). Neurons in the orbitofrontal cortex encode economic value. *Nature*, 441(7090), 223-226.
- Pedroni, A., Langer, N., Koenig, T., Allemand, M., & Jancke, L. (2011). Electroencephalographic Topography Measures of Experienced Utility. *Journal of Neuroscience*, 31(29), 10474-10480. doi: Doi 10.1523/Jneurosci.5488-10.2011
- Perrin, F., Pernier, J., Bertrand, O., & Echallier, J. F. (1989). Spherical Splines for Scalp Potential and Current-Density Mapping. *Electroencephalography and Clinical Neurophysiology*, 72(2), 184-187. doi: Doi 10.1016/0013-4694(89)90180-6
- Pessiglione, M., Seymour, B., Flandin, G., Dolan, R. J., & Frith, C. D. (2006). Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature*, 442(7106), 1042-1045. doi: Doi 10.1038/Nature05051
- Pfabigan, D. M., Alexopoulos, J., Bauer, H., & Sailer, U. (2011). Manipulation of feedback expectancy and valence induces negative and positive reward prediction error signals manifest in event-related brain potentials. *Psychophysiology*, 48(5), 656-664. doi: DOI 10.1111/j.1469-8986.2010.01136.x
- Philiastides, M. G., Biele, G., Vavatzanidis, N., Kazzer, P., & Heekeren, H. R. (2010). Temporal dynamics of prediction error processing during reward-based decision making. *Neuroimage*, 53(1), 221-232. doi: DOI 10.1016/j.neuroimage.2010.05.052
- Pine, A., Seymour, B., Roiser, J. P., Bossaerts, P., Friston, K. J., Curran, H. V., & Dolan, R. J. (2009). Encoding of Marginal Utility across Time in the Human Brain.

Journal of Neuroscience, 29(30), 9575-9581. doi: Doi 10.1523/Jneurosci.1126-09.2009

Potts, G. F., Martin, L. E., Burton, P., & Montague, P. R. (2006). When things are better or worse than expected: The medial frontal cortex and the allocation of processing resources. *Journal of Cognitive Neuroscience*, 18(7), 1112-1119.

Quilodran, R., Rothe, M., & Procyk, E. (2008). Behavioral shifts and action valuation in the anterior cingulate cortex. *Neuron*, 57(2), 314-325. doi: DOI 10.1016/j.neuron.2007.11.031

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. *Classical conditioning II: Current research and theory*, 2, 64-99.

Roberts, K., Infantolino, Z., Stanley, E. M., & Simons, R. F. (unpublished data). University of Delaware.

Rosenthal, R. (1991). *Meta-analytic procedures for social research (rev. ed.)*: Sage Publications, Inc, Thousand Oaks, CA.

Rutledge, R. B., Dean, M., Caplin, A., & Glimcher, P. W. (2010). Testing the Reward Prediction Error Hypothesis with an Axiomatic Model. *Journal of Neuroscience*, 30(40), 13525-13536. doi: Doi 10.1523/Jneurosci.1747-10.2010

Sambrook, T. D., & Goslin, J. (2014). Mediofrontal event-related potentials in response to positive, negative and unsigned prediction errors. *Neuropsychologia*, 61, 1-10. doi: DOI 10.1016/j.neuropsychologia.2014.06.004

Sambrook, T. D., & Goslin, J. (2015). A Neural Reward Prediction Error Revealed by a Meta-Analysis of ERPs Using Great Grand Averages. *Psychological Bulletin*, 141(1), 213-235. doi: Doi 10.1037/Bul0000006

Sambrook, T. D., & Goslin, J. (under revision). Principal components analysis of reward prediction errors in a reinforcement learning task. *Neuroimage*.

- Sambrook, T. D., & Goslin, J. (unpublished data experiment 1). University of Plymouth.
- Sambrook, T. D., & Goslin, J. (unpublished data experiment 2). University of Plymouth.
- Sambrook, T. D., & Goslin, J. (unpublished data experiment 3). University of Plymouth.
- Sambrook, T. D., Roser, M., & Goslin, J. (2012). Prospect theory does not describe the feedback - related negativity value function. *Psychophysiology*, 49(12), 1533-1544.
- San Martin, R. (2012). Event-related potential studies of outcome processing and feedback-guided learning. *Frontiers in Human Neuroscience*, 6. doi: Artn 304
Doi 10.3389/Fnhum.2012.00304
- San Martín, R., Appelbaum, L. G., Pearson, J. M., Huettel, S. A., & Woldorff, M. G. (2013). Rapid Brain Responses Independently Predict Gain Maximization and Loss Minimization during Economic Decision Making. *The Journal of Neuroscience*, 33(16), 7011-7019.
- San Martin, R., Manes, F., Hurtado, E., Isla, P., & Ibanez, A. (2010). Size and probability of rewards modulate the feedback error-related negativity associated with wins but not losses in a monetarily rewarded gambling task. *Neuroimage*, 51(3), 1194-1204. doi: DOI 10.1016/j.neuroimage.2010.03.031
- Santesso, D. L., Dzyundzyak, A., & Segalowitz, S. J. (2011). Age, sex and individual differences in punishment sensitivity: Factors influencing the feedback-related negativity. *Psychophysiology*, 48(11), 1481-1488. doi: DOI 10.1111/j.1469-8986.2011.01229.x
- Sato, A., Yasuda, A., Ohira, H., Miyawaki, K., Nishikawa, M., Kumano, H., & Kuboki, T. (2005). Effects of value and reward magnitude on feedback negativity and P300. *NeuroReport: For Rapid Communication of Neuroscience Research*, 16(4), 407-411. doi: <http://dx.doi.org/10.1097/00001756-200503150-00020>

- Schuermann, B., Endrass, T., & Kathmann, N. (2012). Neural correlates of feedback processing in decision-making under risk. *Frontiers in Human Neuroscience*, 6. doi: Artn 204
- Doi 10.3389/Fnhum.2012.00204
- Schultz, W. (2009). Midbrain dopamine neurons: A retina of the reward system. *Neuroeconomics: Decision making and the brain*, 321-328.
- Schultz, W. (2010). Dopamine signals for reward value and risk: basic and recent data. *Behavioral and Brain Functions*, 6. doi: Artn 24
- Doi 10.1186/1744-9081-6-24
- Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, 275(5306), 1593-1599. doi: DOI 10.1126/science.275.5306.1593
- Seymour, B., Daw, N., Dayan, P., Singer, T., & Dolan, R. (2007). Differential encoding of losses and gains in the human striatum. *Journal of Neuroscience*, 27(18), 4826-4831. doi: Doi 10.1523/Jneurosci.0400-07.2007
- Sutton, R. S., & Barto, A. G. (1998). *Introduction to reinforcement learning*: MIT Press.
- Talmi, D., Atkinson, R., & El-Deredy, W. (2013). The Feedback-Related Negativity Signals Salience Prediction Errors, Not Reward Prediction Errors. *Journal of Neuroscience*, 33(19), 8264-8269. doi: Doi 10.1523/Jneurosci.5695-12.2013
- Talmi, D., Fuentemilla, L., Litvak, V., Duzel, E., & Dolan, R. J. (2012). An MEG signature corresponding to an axiomatic model of reward prediction error. *Neuroimage*, 59(1), 635-645. doi: DOI 10.1016/j.neuroimage.2011.06.051
- Tobler, P. N., Fiorillo, C. D., & Schultz, W. (2005). Adaptive coding of reward value by dopamine neurons. *Science*, 307(5715), 1642-1645. doi: DOI 10.1126/science.1105370

- Tom, S. M., Fox, C. R., Trepel, C., & Poldrack, R. A. (2007). The neural basis of loss aversion in decision-making under risk. *Science*, 315(5811), 515-518. doi: DOI 10.1126/science.1134239
- Toyomaki, A., & Murohashi, H. (2005). Discrepancy between feedback negativity and subjective evaluation in gambling. *Neuroreport*, 16(16), 1865-1868. doi: DOI 10.1097/01.wnr.0000185962.96217.36
- Trepel, C., Fox, C. R., & Poldrack, R. A. (2005). Prospect theory on the brain? Toward a cognitive neuroscience of decision under risk. *Cognitive Brain Research*, 23(1), 34-50. doi: DOI 10.1016/j.cogbrainres.2005.01.016
- Tversky, A., & Kahneman, D. (1981). The Framing of Decisions and the Psychology of Choice. *Science*, 211(4481), 453-458. doi: DOI 10.1126/science.7455683
- Tversky, A., & Kahneman, D. (1992). Advances in Prospect-Theory - Cumulative Representation of Uncertainty. *Journal of Risk and Uncertainty*, 5(4), 297-323. doi: Doi 10.1007/Bf00122574
- van Boxtel, G. J. M., & Bocker, K. B. E. (2004). Cortical measures of anticipation. *Journal of Psychophysiology*, 18(2-3), 61-76. doi: Doi 10.1027/0269-8803.18.2-3.61
- Van den Berg, I., Franken, I. H. A., & Muris, P. (2011). Individual Differences in Sensitivity to Reward Association With Electrophysiological Responses to Monetary Gains and Losses. *Journal of Psychophysiology*, 25(2), 81-86. doi: Doi 10.1027/0269-8803/A000032
- van der Helden, J., Boksem, M. A. S., & Blom, J. H. G. (2010). The Importance of Failure: Feedback-Related Negativity Predicts Motor Learning Efficiency. *Cerebral Cortex*, 20(7), 1596-1603. doi: DOI 10.1093/cercor/bhp224
- Viechtbauer, W. (2010). Conducting Meta-Analyses in R with the metafor Package. *Journal of Statistical Software*, 36(3), 1-48.

- von Neumann, J., & Morgenstern, O. (1944). *Theory of Games and Economic Behaviour*. Princeton, NJ: Princeton University Press.
- Walsh, M. M., & Anderson, J. (unpublished data experiment 1). Carnegie Mellon University
- Walsh, M. M., & Anderson, J. (unpublished data experiment 2). Carnegie Mellon University.
- Walsh, M. M., & Anderson, J. R. (2011a). Learning from delayed feedback: neural responses in temporal credit assignment. *Cognitive Affective & Behavioral Neuroscience*, *11*(2), 131-143. doi: DOI 10.3758/s13415-011-0027-0
- Walsh, M. M., & Anderson, J. R. (2011b). Modulation of the feedback-related negativity by instruction and experience. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(47), 19048-19053. doi: DOI 10.1073/pnas.1117189108
- Walsh, M. M., & Anderson, J. R. (2012). Learning from experience: Event-related potential correlates of reward processing, neural adaptation, and behavioral choice. *Neuroscience and Biobehavioral Reviews*, *36*(8), 1870-1884. doi: DOI 10.1016/j.neubiorev.2012.05.008
- Walsh, M. M., & Anderson, J. R. (2013). Electrophysiological Responses to Feedback during the Application of Abstract Rules. *Journal of Cognitive Neuroscience*, *25*(11), 1986-2002. doi: Doi 10.1162/Jocn_a_00454
- Walsh, M. M., & Anderson, J. R. (2014). Navigating Complex Decision Spaces: Problems and Paradigms in Sequential Choice. *Psychological Bulletin*, *140*(2), 466-486. doi: Doi 10.1037/A0033455
- Weber, B., Aholt, A., Neuhaus, C., Trautner, P., Elger, C. E., & Teichert, T. (2007). Neural evidence for Reference-dependence in real-market-transactions. *Neuroimage*, *35*(1), 441-447. doi: DOI 10.1016/j.neuroimage.2006.11.034

- Weinberg, A., Luhmann, C. C., Bress, J. N., & Hajcak, G. (2012). Better late than never? The effect of feedback delay on ERP indices of reward processing. *Cognitive Affective & Behavioral Neuroscience*, 12(4), 671-677. doi: DOI 10.3758/s13415-012-0104-z
- Wu, Y., & Zhou, X. (2009). The P300 and reward valence, magnitude, and expectancy in outcome evaluation. *Brain Research*, 1286, 114-122. doi: <http://dx.doi.org/10.1016/j.brainres.2009.06.032>
- Wunderlich, K., Smittenaar, P., & Dolan, R. J. (2012). Dopamine Enhances Model-Based over Model-Free Choice Behavior. *Neuron*, 75(3), 418-424. doi: DOI 10.1016/j.neuron.2012.03.042
- Yacubian, J., Glascher, J., Schroeder, K., Sommer, T., Braus, D. F., & Buchel, C. (2006). Dissociable systems for gain- and loss-related value predictions and errors of prediction in the human brain (vol 26, pg 9530, 2006). *Journal of Neuroscience*, 26(39).
- Yasuda, A., Sato, A., Miyawaki, K., Kumano, H., & Kuboki, T. (2004). Error-related negativity reflects detection of negative reward prediction error. *Neuroreport*, 15(16), 2561-2565. doi: Doi 10.1097/00001756-200411150-00027
- Yeung, N., Holroyd, C. B., & Cohen, J. D. (2005). ERP correlates of feedback and reward processing in the presence and absence of response choice. *Cerebral Cortex*, 15(5), 535-544. doi: DOI 10.1093/cercor/bhh153
- Yeung, N., & Sanfey, A. G. (2004). Independent coding of reward magnitude and valence in the human brain. *Journal of Neuroscience*, 24(28), 6258-6264. doi: Doi 10.1523/Jneurosci.4537-03.2004
- Yi, F., Chen, H., Wang, X., Shi, H., Yi, J., Zhu, X., & Yao, S. (2012). Amplitude and latency of feedback-related negativity: Aging and sex differences. *NeuroReport*:

For Rapid Communication of Neuroscience Research, 23(16), 963-969. doi:

<http://dx.doi.org/10.1097/WNR.0b013e328359d1c4>

Yu, R. J., Zhou, W., & Zhou, X. L. (2011). Rapid Processing of Both Reward

Probability and Reward Uncertainty in the Human Anterior Cingulate Cortex.

Plos One, 6(12). doi: ARTN e29633DOI 10.1371/journal.pone.0029633

Yu, R. J., & Zhou, X. L. (2006). Brain responses to outcomes of one's own and other's

performance in a gambling task. *Neuroreport*, 17(16), 1747-1751. doi: DOI

10.1097/01.wnr.0000239960.98813.50

Yu, R. J., & Zhou, X. L. (2009). To Bet or Not to Bet? The Error Negativity or Error-

related Negativity Associated with Risk-taking Choices. *Journal of Cognitive*

Neuroscience, 21(4), 684-696. doi: DOI 10.1162/jocn.2009.21034

Zhong, S. F., Israel, S., Xue, H., Sham, P. C., Ebstein, R. P., & Chew, S. H. (2009). A

neurochemical approach to valuation sensitivity over gains and losses.

Proceedings of the Royal Society B-Biological Sciences, 276(1676), 4181-4188.

doi: DOI 10.1098/rspb.2009.1312

Zottoli, T. M., & Grose-Fifer, J. (2012). The feedback-related negativity (FRN) in

adolescents. *Psychophysiology*, 49(3), 413-420. doi: DOI 10.1111/j.1469-

8986.2011.01312.x